This page intentionally left blank.

Best viewed with double-page viewing.

Hermann G. Matthies

A Hitchiker's Guide to Notation and Definitions

Informatikbericht Nr. 2011-05

Institute of Scientific Computing
Carl-Friedrich-Gauss-Fakultät
Technische Universität Braunschweig

Brunswick, Germany

This document was created July 2011 using $\LaTeX\,2_\varepsilon$.

Institute of Scientific Computing
Technische Universität Braunschweig
Hans-Sommer-Straße 65
D-38106 Braunschweig, Germany

url: www.wire.tu-bs.de
mail: wire@tu-bs.de

A Hitchiker's Guide to Notation and Definitions

Hermann G. Matthies

Institute of Scientific Computing, TU Braunschweig
wire@tu-bs.de

Abstract

This little note is intended to make some suggestions for a consistent and on the other hand mostly accepted notation. This is deemed to be an aid when reading mathematics, as it can be clear from the denotation of an object which type it is. With allusions to D. Adams [4], and following [5], this is intended to be a useful compilation. The main point though is:

Don't panic!

Keywords: logic and sets, linear algebra and analysis, topology, measure spaces, functional analysis, manifolds, convex analysis, stochastic analysis

Contents

1 General Notation

As most of this is going to be on vector spaces, vectors, operators, topologies, σ-algebras, and the like, here some suggestions for notation (obviously this is not to be used if there is some generally recognised standard notation).

The generally used notation for elementary analysis is collected in this part. More specialised subjects will follow in later sections. Observe also [1].

Whereas "=" is the *equal sign*—historically chosen because the two lines of the sign are equal—the notation "$A := B$" means that A is *defined* to be B, and similarly for $B =: A$.

One way to differentiate different quantities is through the actual notation, some of which is collected in section 9, and through the use of different fonts collected in section 14. This is to allow to recognise the type of a mathematical object from its denotation.

1.1 Logic

If $\mathfrak{a}$ and $\mathfrak{b}$ are (logical) statements, then $\mathfrak{a} \Leftrightarrow \mathfrak{b}$ is their logical **equivalence**, $\neg\mathfrak{a}$ is the **negation**, $\mathfrak{a} \wedge \mathfrak{b}$ is their **conjunction (logical and)**, and $\mathfrak{a} \vee \mathfrak{b}$ is their **disjunction (logical or)**, and $\mathfrak{a} \Rightarrow \mathfrak{b}$ is the logical **implication** $\neg\mathfrak{a} \vee \mathfrak{b}$.

If $\{\mathfrak{a}_j\}_{j\in J}$ is a collection or family of statements, their **total conjunction** is denoted $\forall j \in J : \mathfrak{a}_j$ or $\bigwedge_{j\in J} \mathfrak{a}_j$, and their **total disjunction** by $\exists j \in J : \mathfrak{a}_j$ or $\bigvee_{j\in J} \mathfrak{a}_j$.

1.2 Sets

Usually sets are denoted by upper case, like A, B, M, and elements by lower case letters, i.e. $a \in A$. The empty set is denoted as $\emptyset$. Let A and B be **subsets** of some set M, $A \subseteq M$ and $B \subseteq M$. This notation is understood to mean that A and B may possibly be all of M, whereas are $A \subset M$ then means that A is a *proper* subset, i.e. $A \subset M \Leftrightarrow (A \subseteq M) \wedge (A \neq M)$.

For such sets, $A \cup B$ denotes their **union**, $A \cap B$ the **intersection**. The **complement** of a set is denoted by $\complement A := A^c := \{a : a \in M \wedge \neg(a \in A)\}$. The **set difference** is $A \setminus B := A \cap \complement B$, and the **symmetric set difference** is $A \bigtriangleup B := (A \setminus B) \cup (B \setminus A)$. Observe that $A \bigtriangleup \emptyset = A$ for any set A. For two sets A and B, their **Cartesian product** is $A \times B$, elements are *ordered tuples* $(a, b) \in A \times B$ with $a \in A$ and $b \in B$. The **disjoint union** of two disjoint sets A and B with $A \cap B = \emptyset$ is just $A \uplus B := A \cup B$. In case $A \cap B \neq \emptyset$, one may define $A \uplus B := \{(1, a) : a \in A\} \cup \{(2, b) : b \in B\}$, a construction which makes the sets artificially disjoint.

If $\mathfrak{C} = \{A_j\}_{j\in J}$ is a collection or family of sets with index set J, their *union* is denoted $\bigcup_{j\in J} A_j = \bigcup_{B\in\mathfrak{C}} B$, and their *Cartesian product* is $\prod_{j\in J} A_j$, for a finite number n of sets also often denoted by $A_1 \times A_2 \times \cdots \times A_n$. An element of this *Cartesian product* is an *ordered tuple* $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$ with $a_j \in A_j$. If the sets A_j in a family are all equal to some set A, $A^J := \prod_{j\in J} A$ may be seen as the set of all mappings from J into A. For a finite number $n \in \mathbb{N}$ of sets, this is mostly denoted as A^n, as $n = |\{1, \ldots, n\}|$. The **power-set** of A is denoted by 2^A, the set of all maps from A into $\{0, 1\}$, with $|\{0, 1\}| = 2$. Observe that $2^{\emptyset} = \{\emptyset\}$.

The **disjoint union** of a collection $\mathfrak{C} = \{A_j\}_{j\in J}$ is denoted $\biguplus_{j\in J} A_j$, or also $\coprod_{j\in J} A_j$, or sometimes $\sum_{j\in J} A_j$, which is just the union if the sets A_j are mutually disjoint. In case the sets in the family are not mutually disjoint, they can easily be made so as just explained by defining for each j the new sets $\tilde{A}_j := \{j\} \times A_j$, and then using the $\tilde{A}_j$ instead of the original A_j, i.e. $\biguplus_{j\in J} A_j \equiv \bigcup_{j\in J} \tilde{A}_j$, identified under the mapping $\forall j \in J : A_j \ni a \mapsto (j, a) \in \tilde{A}_j$. If all A_j are equal to some set A, then $\biguplus_{j\in J} A_j \equiv J \times A$.

The **cardinality** (number of elements) is written $|A| := \mathrm{card}(A)$. Observe that $|A \times B| = |A| \cdot |B|$, $|A^n| = n|A|$, $|2^A| = 2^{|A|}$, and $|A \uplus B| = |A| + |B|$. Some special cases are $|\emptyset| = 0$, $|\{\emptyset\}| = 1$, with notations $|\mathbb{N}| =: \aleph_0$ (the Hebrew letter *aleph*: $\aleph$) and $|\mathbb{R}| =: \mathfrak{c}$. Remember that $\forall n \in \mathbb{N} : 0 < n < \aleph_0$ and $\aleph_0 < \mathfrak{c}$. A set with $|A| < \aleph_0$ is called **finite**, otherwise **infinite**. If $|A| \leq \aleph_0$ it is called **countable**, otherwise **uncountable**.

If all elements A_j of the collection $\mathfrak{C}$ are subsets of some set M, i.e. $\mathfrak{C} \subseteq 2^M$, their *intersection* is denoted by $\bigcap_{j\in J} A_j$. In case $\forall i, j \in J : \quad i \neq j \Rightarrow A_i \cap A_j = \emptyset$, the collection is **mutually disjoint**. If $\bigcap_{j\in J} A_j = \emptyset$, the collection of sets is called **free**, otherwise it is **fixed**. In case every finite subfamily $\{A_{j^1}, \ldots, A_{j^n}\} \subseteq \mathfrak{C}$ satisfies $\bigcap_{k=1}^{n} A_{j^k} \neq \emptyset$, one says that the collection $\mathfrak{C}$ has the **finite intersection property**.

Such a nonempty collection $\mathfrak{C} \neq \emptyset$ is called a **covering** of $B \subseteq M$, if $\bigcup_{j\in J} A_j \supseteq B$. If $\mathfrak{B} = \{B_k\}_{k\in K}$ is another collection, such that $\bigcup_{k\in K} B_k \supseteq B$ is another covering, it is called a **refinement** if for each $k \in K$ there is a $j \in J$ such that $B_k \subseteq A_j$. A covering is called **point-finite** if for any $a \in B$ only for finitely many A_j one has $a \in A_j$. A **partition** of B is a covering $\bigcup_{j\in J} A_j = B$ with mutually disjoint nonempty subsets $\forall j \in J : \emptyset \neq A_j \subseteq B$.

Observe that when $J = \emptyset$, then $\{A_j\}_{j\in J} = \mathfrak{C} = \emptyset \subset 2^M$, and $\bigcup_{j\in J} A_j = \emptyset$, $\prod_{j\in J} A_j = \emptyset$, whereas $\bigcap_{j\in J} A_j = M$. Also observe that if any of the $A_j = \emptyset$, this implies $\prod_{j\in J} A_j = \emptyset$, and if $J \neq \emptyset$ also $\bigcap_{j\in J} A_j = \emptyset$.

In this connection, remember the **Axiom of Choice**, which says that for a nonempty collection of nonempty sets, also $\prod_{j\in J} A_j$ is non-empty; i.e. there is an element $a = (\ldots, a_j, \ldots) \in \prod_{j\in J} A_j$, so that from each A_j one may *choose* an element a_j.

Connected with a union or disjoint union of a family $\{A_j\}_{j\in J}$ are the **canonical injections** $\forall k \in J : \iota_k : A_k \to \bigcup_{j\in J} A_j$, resp. $\iota_k : A_k \to \biguplus_{j\in J} A_j = \coprod_{j\in J} A_j$, such that $\iota_k : A_k \ni a_k \mapsto a_k \in \bigcup_{j\in J} A_j$. Similarly, with an intersection $\bigcap_{j\in J} A_j$ associate the **canonical "projections"** $\forall k \in J : \rho_k : \bigcap_{j\in J} A_j \to A_k$, such that $\rho_k : \bigcap_{j\in J} A_j \ni a \mapsto a \in A_k$. These are essentially the identity with different *domain* or *co-domain*—cf. section 1.6. Associated with the product are the "real" **canonical projections** $\forall k \in J : \pi_k : \prod_{j\in J} A_j \to A_k$ such that $\pi_k : \prod_{j\in J} A_j \ni a = (a_1, \ldots, a_k, \ldots) \mapsto a_k \in A_k$.

If G is an **Abelian** group or monoid written additively with neutral element 0, the **direct sum** $G^{(J)} = \bigoplus_{j\in J} G \subseteq G^J$ is the group or monoid of sequences $(g_j)_{j\in J}$, where only finitely many $g_j \neq 0$ (this is a special case of the categorical construction of co-product). If $|J| < \infty$, both constructions (product and sum) essentially coincide. With the direct sum one may again associate the **canonical injections** $\forall k \in J : \iota_k : G_k \to \bigoplus_{j\in J} G_j$.

Other than the union or intersection of a family of sets, there are two other useful constructions. If J is a directed set (cf. Section 2.1), the **limes superior** of a net $\{A_j\}_{j\in J} \subseteq 2^M$ of subsets of M is:

$$\limsup_{j\in J} A_j := \bigcap_{j\in J} \bigcup_{j\preceq k} A_k \subseteq M.$$

If $|J| \geq \aleph_0$, the limes superior are those elements of M which are contained in infinitely many A_j. Similarly, one may define the **limes inferior** via

$$\liminf_{j\in J} A_j := \bigcup_{j\in J} \bigcap_{j\preceq k} A_k \subseteq M.$$

Again if $|J| \geq \aleph_0$, the limes inferior are those elements of M which are contained in all but finitely many A_j. Obviously $\liminf_{j\in J} A_j \subseteq \limsup_{j\in J} A_j$, and if both of them are equal, the net is called **convergent**, with the limit

$$\lim_{j\in J} A_j := \liminf_{j\in J} A_j = \limsup_{j\in J} A_j \subseteq M.$$

A **filter** $\mathfrak{F} \subseteq 2^M$ is a nonempty collection of subsets of a set M such that $\mathfrak{F}$ is closed under finite intersections and contains all supersets of its members, i.e. $S, T \in \mathfrak{F} \Rightarrow S \cap T \in \mathfrak{F}$, and $(S \subseteq T \subseteq M \wedge S \in \mathfrak{F}) \Rightarrow T \in \mathfrak{F}$. In case that $\mathfrak{F}$ does not contain the empty set (otherwise $\mathfrak{F} = 2^M$), it is called a **proper filter**. A **filter basis** $\mathfrak{B} \subseteq 2^M$ is a nonempty collection which does not contain the empty set, and such that for $S, T \in \mathfrak{B} \Rightarrow \exists U \in \mathfrak{B} : U \subseteq S \cap T$.

Then $\mathfrak{F}(\mathfrak{B}) = \{S \subseteq M | \exists T \in \mathfrak{B} : T \subseteq S\}$ is the *proper* filter generated by $\mathfrak{B}$. This is the *smallest* filter—with regards to inclusion—which contains $\mathfrak{B}$; equivalently, this is the intersection of all such filters.

An important example for a filter is $\mathfrak{F} = \{S \subseteq M : \complement S \text{ is finite}\}$, the **Fréchet filter** of **co-finite** sets, where M is nonempty. It is proper iff $|M| \geq \aleph_0$. For a net or sequence $\{a_j\}_{j \in J}$ in M, the collection $\mathfrak{B} = \{\{a_\ell\}_{\ell \in L_k} | \, L_k = \{\ell \in J : k \preceq \ell\}, k \in J\}$ is a filter basis for the **tail filter** of the net or sequence.

In case $\mathfrak{F}$ is a proper filter on M, and $\forall S \subseteq M$, either $S \in \mathfrak{F}$ or $\complement S \in \mathfrak{F}$, then $\mathfrak{F}$ is called an **ultrafilter.** Equivalently, $\mathfrak{F}$ is a *maximal* filter (see section 1.5). Note that if the ultrafilter is free and the set M is infinite, then an ultrafilter $\mathfrak{F}$ contains no finite set, and is a superset of the Fréchet filter.

1.3 Relations

A (binary) relation $\mathcal{R}$ is a subset of $A \times B$. For $(a, b) \in \mathcal{R} \subseteq A \times B$ one also writes $a \, \mathcal{R} \, b$ or $\mathcal{R}(a, b)$. A (binary) relation on a set A is a subset $\mathcal{R} \subseteq A^2$, and is called:

reflexive if $(a, a) \in \mathcal{R}$, i.e. if $a \, \mathcal{R} \, a$.

irreflexive if $\neg \, (a \, \mathcal{R} \, a)$.

symmetric if $a \, \mathcal{R} \, b \;\Rightarrow\; b \, \mathcal{R} \, a$.

asymmetric if $a \, \mathcal{R} \, b \;\Rightarrow\; \neg(b \, \mathcal{R} \, a)$.

antisymmetric if $(a \, \mathcal{R} \, b) \wedge (b \, \mathcal{R} \, a) \Rightarrow (a = b)$.

transitive if $(a \, \mathcal{R} \, b) \wedge (b \, \mathcal{R} \, c) \Rightarrow (a \, \mathcal{R} \, c)$.

total if $(a \neq b) \;\Rightarrow\; (a \, \mathcal{R} \, b) \vee (b \, \mathcal{R} \, a)$.

complete if $(a \, \mathcal{R} \, b) \vee (b \, \mathcal{R} \, a)$.

Observe that a complete relation is total. The **inverse relation** is denoted by $(a \, \mathcal{R}^{-1} \, b) \Leftrightarrow (b \, \mathcal{R} \, a)$. The **composition** of relations $\mathcal{R}$ and $\mathcal{S}$ on a set A is denoted by $(a \, \mathcal{S} \circ \mathcal{R} \, b) \Leftrightarrow \exists c \in A : \; (c \, \mathcal{S} \, b) \wedge (a \, \mathcal{R} \, c)$. If $\mathcal{R}$ is composed with itself, this is denoted by $\mathcal{R}^2 = \mathcal{R} \circ \mathcal{R}$; similarly for higher powers.

1.4 Equivalence Relations

An equivalence relation is a *reflexive, symmetric, transitive* relation. It is usually denoted similarly to $a \sim b := a \mathcal{R} b$ or $a \simeq b := a \mathcal{R} b$. For an equivalence relation on a set A, the **equivalence class** $[a]$ of an element $a \in A$ is $[a] := \{b \in A : b \sim a\} \subseteq A$. The equivalence classes are a partition of the set A (section 1.2). The set of equivalence classes is called the **quotient** of A w.r.t. the equivalence relation $\sim$, and is denoted by $A/\sim$.

In turn, any partition $\{A_j\}_{j \in J}$ of a set A induces an equivalence relation via $a \sim b \Leftrightarrow \exists j \in J : a, b \in A_j$.

With any equivalence relation $\sim$ one may associate the **canonical projection** $q : A \to A/\sim$ with $q(a) = [a]$.

1.5 Order Relations

Order relations come in different flavours, they are written in a manner similarly to $a \preceq b := a \mathcal{R} b$, and what is usually meant is a *partial order*. A relation is an *order relation* if it one of the following:

pre-order if it is *reflexive, transitive.*

partial order if it is an *antisymmetric* pre-order, i.e. it is *reflexive, antisymmetric, transitive.*

linear order if it is a *total* or *complete* partial order.

strict partial order (denoted similarly to $a \prec b := a \mathcal{R} b$) if it is *asymmetric, irreflexive, transitive.*

strict linear order if it is a *total* strict partial order.

In the context of a *pre-order* or a *partial order*, $a \prec b$ is often taken to mean $(a \preceq b) \wedge (a \neq b)$. A set A together with a partial order is called a partially ordered set, or sometimes a **poset**.

An element $a \in B$ is a **maximal** element of B, if $\forall b \in B : (a \preceq b) \Rightarrow (a = b)$, i.e. if there is no other element $b \in B, b \neq a$ with $a \preceq b$. Similarly for a **minimal** element.

An **upper bound** for a subset $B \subseteq A$ is an element $a \in A$ with $\forall b \in B : b \preceq a$. Similarly for a **lower bound**. An **order interval** is $[a, b] := \{a' \in A : a \preceq a' \preceq b\} \subseteq A$, and similarly $]a, b[:= \{a' \in A : a \prec a' \prec b\} \subseteq A$. A subset B is **order bounded** iff B is contained in an order interval $[a, b]$, i.e. if there is both a lower and an upper bound.

An element $a \in B$ is a **greatest** or a **maximum** element of B —denoted $\max B$ —if $\forall b \in B : b \preceq a$, i.e. a is an upper bound which is an element of the set. Similarly for a **least** or **minimum** element, denoted by $\min B$. The **greatest lower bound** is called the **infimum** $\inf B$, and the **supremum** $\sup B$ is the **least upper bound.** For the two-element set $\{a, b\}$ one writes $a \vee b := \sup\{a, b\}$, and $a \wedge b := \inf\{a, b\}$. For linear orders $a \vee b = \max\{a, b\}$ and $a \wedge b = \min\{a, b\}$.

A **directed set** is a set with a **pre-order** where every pair of elements has an upper bound. A **chain** is a linearly ordered nonempty subset of a poset. A poset where every chain has a supremum is **inductively** ordered. A poset is **well ordered** if every nonempty subset has a minimum. A **lattice** is a poset where every pair of elements has a supremum and an infimum.

In this connection, remember **Zorn's lemma**, which says that in an *inductively ordered* poset there is a maximal element. This is equivalent to the *axiom of choice*, see section 1.2.

Also, observe that according to the **well ordering principle**, any set may be well ordered. This is equivalent to *Zorn's lemma.*

If A and B are posets, a function $\phi : A \to B$ is **isotone** (sometimes called **monotone**) if $a \preceq b$ on A implies $\phi(a) \preceq \phi(b)$ on B.

1.6 Maps

A **map**, **mapping** or a **function** ϕ from A to B assigns to each $a \in A$ one element $b = \phi(a) \in B$. The set A is called the **domain**, and B the **co-domain**. One usually writes $\phi : A \to B$, or $\phi \in (A \to B)$, or $\phi \in B^A$, or $\phi \in \mathscr{F}(A, B)$, where each of the notations means the set of all maps from A into B. In particular $b = \phi(a)$ may be also denoted $\phi : a \mapsto b$. The mapping alone is just denoted by ϕ or with more emphasis as $\phi(\cdot)$, but often $\phi(a)$ may mean the whole mapping instead of just the single value $\phi(a) \in B$. The intended use is normally clear from the context.

For a subset $C \subseteq A$ and a mapping $\phi \in \mathscr{F}(A, B)$, the restriction of ϕ to C is denoted by $\phi|_C \in \mathscr{F}(C, B)$. The set $\phi(A) =: \operatorname{im} \phi \subseteq B$ is the **image** or **range** of ϕ. In case ϕ is only a **partial** function (not defined on all of A), the set $\operatorname{dom} \phi \subset A$ is the **domain** of its definition. For $E \subseteq B$, the **pre-image** or **inverse image** of E is $\phi^{-1}(E) = \{a \in A : \phi(a) \in E\}$. The **graph** of a map is

$$\operatorname{gra} \phi := \{(a, b) \in \operatorname{dom} \phi \times \operatorname{im} \phi |\, b = \phi(a)\} \subseteq A \times B.$$

Maps where $\operatorname{im} \phi \subseteq \mathbb{K}$ are often called **functionals**. Some special functions in $\mathscr{F}(A, \mathbb{R})$ are the **characteristic function** of a subset $B \subseteq A$ (in *probability theory* this function is often called the indicator, conflicting with the definition below, and the characteristic function is the *Fourier transform* of a random variable),

$$\forall a \in A: \ \chi_B(a) := \begin{cases} 1, & \text{if } a \in B \\ 0, & \text{otherwise,} \end{cases}$$

the **Kronecker symbol** $\forall a, b \in A: \ \delta_{ab} := \chi_{\{a\}}(b) = \chi_{\{b\}}(a)$, sometimes also denoted by δ_a^b, and the **indicator function** of a subset $B \subseteq A$,

$$\forall a \in A: \ \psi_B(a) := \begin{cases} 0, & \text{if } a \in B \\ +\infty, & \text{otherwise.} \end{cases}$$

In this case one calls $\operatorname{dom} \psi_B := B$ the **effective** domain.

Another special map in $\mathscr{F}(A, A)$ is the identity mapping, having various notations, e.g. I, $\boldsymbol{I}$, $\imath$, or ι, or even I_A to identify the set.

A map $\phi \in \mathscr{F}(A, B)$ is **surjective** or **onto** iff $\phi(A) = B$. It is **injective** or **one-to-one** iff $\phi(a_1) = \phi(a_2) \Rightarrow a_1 = a_2$. It is **bijective** if it is both injective and surjective. Note that if a map ϕ is surjective, it has a **right inverse** (necessarily *injective*) $\psi \in \mathscr{F}(B, A)$ with $\phi \circ \psi = I_B$, where I_B is the identity on B. If ϕ is injective, it has a **left inverse** (necessarily *surjective*) $\varphi \in \mathscr{F}(B, A)$ with $\varphi \circ \phi = I_A$. If it is bijective, it has a unique (*bijective*) **inverse** $\phi^{-1} \in \mathscr{F}(B, A)$. Any map $\phi \in \mathscr{F}(A, B)$ induces an equivalence relation on A via $a_1 \sim a_2 \Leftrightarrow \phi(a_1) = \phi(a_2)$. A map $\phi \in \mathscr{F}(A, A)$ is **idempotent** in case $\phi^2 := \phi \circ \phi = \phi$. It is called an **involution** if $\phi^2 = I_A$.

Note that if B is a group, ring, vector space, or an algebra, so is $\mathscr{F}(A, B)$ with the operations defined pointwise. This is especially the case for the functionals $\mathscr{F}(A, \mathbb{K})$, as $\mathbb{K}$ is a *field*.

In case $\varphi \in \mathscr{F}(A, C)$ and $\psi \in \mathscr{F}(B, D)$, this is naturally extended to situations where $A \times C \ni (a, b) \mapsto (\varphi(a), \psi(b)) \in C \times D$, the resulting map is denoted $(\varphi \times \psi) \in \mathscr{F}(A \times B, C \times D)$. Similarly for the situation $A \oplus B$ or $A \otimes B$ in case both A, B, C, D are groups or vector spaces; the resulting maps are denoted by $\varphi \oplus \psi$, or $\varphi \otimes \psi$ respectively, see section 4.2.

One says that $\mathcal{F} \subseteq \mathscr{F}(A, B)$ separates the points in A, if for any $a, b \in A$, $a \neq b$ there is a $\phi \in \mathcal{F}$ such that $\phi(a) \neq \phi(b)$. For such a separating subset of functions, the set A can be identified with a subset of $\mathscr{F}(\mathcal{F}, B)$. Associate to an element $a \in A$ the **evaluation** or **delta**-functional δ_a, which for $\phi \in \mathcal{F}$ is defined as $\delta_a(\phi) := \phi(a)$. The function $a \mapsto \delta_a$ maps A injectively into $\mathscr{F}(\mathcal{F}, B)$. Here A and $\mathcal{F}$ play symmetric rôles, and to emphasise this one may write $\langle \phi, a \rangle := \phi(a) = \delta_a(\phi)$; where $\langle \cdot, \cdot \rangle : \mathcal{F} \times A \to B$ is the **evaluation duality.**

1.7 Special Sets

For the natural numbers, use $\mathbb{N}$, for the integers $\mathbb{Z}$, for the rationals $\mathbb{Q}$, for the reals $\mathbb{R}$, for the complex numbers $\mathbb{C}$, and for a generic field $\mathbb{K}$. For a complex number $z = x + \mathrm{i}y \in \mathbb{C}$, the real part is $x = \Re z$, the imaginary part is $y = \Im z$, and the conjugate complex number is $\bar{z} := z^* := x - \mathrm{i}y$. The imaginary unit "i" is denoted differently from the usual index $\imath, i \in \mathbb{N}$, and also from the notation for the injection map ι. The **absolute value** or **modulus** is $|z| = \sqrt{zz^*}$, and the angle $-\pi < \varphi \leq \pi$ in the representation $z = |z|\,\mathrm{e}^{\mathrm{i}\varphi}$ is the **argument**. Note that the base of natural logarithms "e" is not a variable and is denoted differently.

More specifically, one denotes by $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$ the non-negative integers, the right half-line by $\mathbb{R}_+ := \{r \in \mathbb{R} : r \geq 0\}$, the positive multiplicative group by $\mathbb{R}_* := \{r \in \mathbb{R} : r > 0\}$, the semi-extended reals by $\mathbb{R}^+ := \mathbb{R} \cup \{+\infty\}$, the extended reals by $\overline{\mathbb{R}} := \{-\infty\} \cup \mathbb{R}^+$, the right half-plane by $\mathbb{C}_+ := \{z \in \mathbb{C} | \Re z \geq 0\}$, the left half-plane by $\mathbb{C}_- := \{z \in \mathbb{C} | \Re z \leq 0\}$, the upper half-plane by $\mathbb{C}^+ := \{z \in \mathbb{C} | \Im z \geq 0\}$, the lower half-plane by $\mathbb{C}^- := \{z \in \mathbb{C} | \Re z \geq 0\}$, and the one-point compactification of the complex plane by $\overline{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$.

The set $\{x \in \mathbb{R} \mid a \leq x \leq b\} =: [a, b]$ is a closed interval, open intervals $\{x \in \mathbb{R} \mid a < x < b\}$ are denoted as $]a, b[$, or somewhat ambiguously (a, b). Half-open intervals are correspondingly denoted by $[a, b[$ or $[a, b)$.

Note especially the set (monoid) of all sequences of positive integers $\mathcal{N} = \mathbb{N}^{\mathbb{N}}$, similarly the set (semi-group) $\mathcal{N}_0 := \mathbb{N}_0^{\mathbb{N}}$, where the sequences may include zero, and $\mathcal{N}_{00} := \mathbb{N}_0^{(\mathbb{N})}$, the set (semi-group) of sequences with only *finitely* many non-zeros.

Analogously, $\boldsymbol{\omega} := \mathbb{R}^{\mathbb{N}} = \prod_{k \in \mathbb{N}} \mathbb{R}$ is the vector space of all real valued sequences, and $\boldsymbol{c}_{00} := \mathbb{R}^{(\mathbb{N})} = \bigoplus_{k \in \mathbb{N}} \mathbb{R} \subset \boldsymbol{\omega}$ is the subspace of all finite real sequences $(\varrho_n)_{n \in \mathbb{N}}$, nonzero $(\varrho_n \neq 0)$ only for finitely many $n \in \mathbb{N}$. Of course

$$\forall n \in \mathbb{N} : \mathbb{K}^n = \prod_{k=1}^{n} \mathbb{K} \cong \bigoplus_{k=1}^{n} \mathbb{K}.$$

In this context, one may denote scalars preferably with lower case *Greek* letters, or by lower case *Latin* letters, i.e. $\alpha \in \mathbb{K}$ or $x \in \mathbb{K}$. Subsets may be denoted by upper case *Greek* or *Latin* letters, i.e. $\mathit{\Omega} \subseteq \mathbb{K}$ or $A \subset \mathbb{K}$. The notation $\mathit{\Omega}$ is here preferred to Ω. Usually $\imath, \jmath, k, \ell, m, n \in \mathbb{N}$.

1.8 Matrices

This is an enlargement on recommendations by *Householder*. An element of either $\mathbb{K}^n$, $\boldsymbol{\omega}$, or $\boldsymbol{c}_{00}$ is preferably denoted by a lower case (here bold, to set it apart from an abstract element of a vector space) math Latin letter $\boldsymbol{x} \in \mathbb{K}^n$, realised as a *column* matrix $\boldsymbol{x} \in \mathbb{K}^{n\times 1}$. The **canonical** basis is $\{\boldsymbol{e}_\jmath\} \in \mathbb{K}^{n\times 1}$, where $\boldsymbol{e}_\jmath$ has a 1 in the $\jmath$-th position and 0 otherwise.

Matrices are denoted by an upper case (here bold, to set them apart from abstract [linear] mappings) math letter $\boldsymbol{A} \in \mathbb{K}^{n\times m}$, and $\boldsymbol{A}^T$ is the transpose of $\boldsymbol{A}$, whereas $\boldsymbol{A}^*$ is the conjugate complex transpose of $\boldsymbol{A}$. The unit matrix is denoted by $\boldsymbol{I}$, and the zero matrix (and zero vector) by $\mathbf{0}$.

In linear combinations of matrices the unit matrix is often omitted, i.e. with $\alpha, \beta \in \mathbb{K}$, an expression such as $\alpha\boldsymbol{A} + \beta\boldsymbol{I}$ is sometimes denoted by $\alpha\boldsymbol{A} + \beta$.

The **determinant** of a square matrix $\boldsymbol{A}$ is denoted $\det \boldsymbol{A}$. Note that the determinant is a completely skew-symmetric or alternating n-linear mapping into $\mathbb{K}$ as regards the columns of $\boldsymbol{A} = [\boldsymbol{a}_1, \dots, \boldsymbol{a}_n] \in (\mathbb{K}^n)^n$, and as the space of such maps is one-dimensional, it is completely specified by letting $\det \boldsymbol{I} = 1$.

The matrix product is denoted by $\boldsymbol{AB} = \boldsymbol{A}\cdot\boldsymbol{B}$, so that $\langle \boldsymbol{y}|\boldsymbol{x}\rangle := \boldsymbol{y}^T\boldsymbol{x} = \sum_k y_k x_k$ (or $\langle \boldsymbol{y}|\boldsymbol{x}\rangle := \boldsymbol{y}^*\boldsymbol{x} = \sum_k \bar{y}_k x_k$ in case of complex vectors) is the *canonical inner product* , also denoted by the **dot product** $\boldsymbol{y}\cdot\boldsymbol{x} := \boldsymbol{y}^T\boldsymbol{x}$.

The **outer product** of two vectors $\boldsymbol{x} \in \mathbb{K}^n$ and $\boldsymbol{y} \in \mathbb{K}^m$ is $\boldsymbol{y}\otimes\boldsymbol{x} := \boldsymbol{x}\boldsymbol{y}^T \in \mathbb{K}^{n\times m}$ (note the interchange), also called the **dyadic product**, or *tensor product* (see section 8), here simply seen as a matrix. When applying a matrix $\boldsymbol{A} \in \mathbb{K}^{n\times m}$ to a vector $\boldsymbol{y} \in \mathbb{K}^m$ we write $\boldsymbol{x} = \boldsymbol{A}\boldsymbol{y} \in \mathbb{K}^n$, so that $(\boldsymbol{y}\otimes\boldsymbol{x})\boldsymbol{z} = \langle\boldsymbol{y}|\boldsymbol{z}\rangle\boldsymbol{x} = (\boldsymbol{x}\boldsymbol{y}^T)\boldsymbol{z} = (\boldsymbol{y}\cdot\boldsymbol{z})\boldsymbol{x}$.

The **trace** of a square matrix is denoted by $\operatorname{tr}\boldsymbol{A}$. Note that the trace is linear, i.e. an element of the dual space of $\mathbb{K}^{n\times n}$, completely determined by requiring that $\operatorname{tr}(\boldsymbol{y}\boldsymbol{x}^T) := \boldsymbol{y}^T\boldsymbol{x} = \boldsymbol{y}\cdot\boldsymbol{x}$.

The **tensor product** of two matrices $\boldsymbol{A} = (a_{\imath\jmath}) \in \mathbb{K}^{k\times\ell}$ and $\boldsymbol{B} \in \mathbb{K}^{n\times m}$ (see section 8) is written as $\boldsymbol{A}\otimes\boldsymbol{B}$, and acts on $\boldsymbol{x}\otimes\boldsymbol{y} \in \mathbb{K}^\ell\otimes\mathbb{K}^m$ as $(\boldsymbol{A}\otimes\boldsymbol{B})\boldsymbol{x}\otimes\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{x})\otimes(\boldsymbol{B}\boldsymbol{y}) \in \mathbb{K}^k\otimes\mathbb{K}^n$.

The analogous action is obtained when $\boldsymbol{x}\otimes\boldsymbol{y}$ is interpreted as a matrix $\boldsymbol{y}\boldsymbol{x}^T$ via $(\boldsymbol{A}\otimes\boldsymbol{B})\boldsymbol{x}\otimes\boldsymbol{y} := \boldsymbol{B}\boldsymbol{y}\boldsymbol{x}^T\boldsymbol{A}^T = (\boldsymbol{A}\boldsymbol{x})\otimes(\boldsymbol{B}\boldsymbol{y})$, now interpreted as a matrix $(\boldsymbol{B}\boldsymbol{y})(\boldsymbol{A}\boldsymbol{x})^T$ in $\mathbb{K}^{k\times n}$, isomorphic to $\mathbb{K}^k\otimes\mathbb{K}^n$ and $\mathbb{K}^{k\cdot n}$. Again, it should be clear from the context what is meant.

The **Kronecker product** of two matrices $\boldsymbol{A}$ and $\boldsymbol{B}$ is usually also denoted by $\boldsymbol{A}\otimes\boldsymbol{B} \in \mathbb{K}^{k\cdot n\times\ell\cdot m}$, and is defined as the matrix with block elements $(a_{\imath\jmath}\boldsymbol{B})$. It can act on a dyadic or tensor product of two vectors $\boldsymbol{x} \in \mathbb{K}^\ell$ and $\boldsymbol{y} \in \mathbb{K}^m$ if this is also consistently interpreted as a Kronecker product of matrices $\boldsymbol{x}\otimes\boldsymbol{y} := [x_1\boldsymbol{y}^T, \dots, x_\ell\boldsymbol{y}^T]^T \in \mathbb{K}^{\ell\cdot m}$. The result is obviously a vector in $\mathbb{K}^{k\cdot n}$, which is consistent with the tensor product interpretation with the canonical isomorphism between $\mathbb{K}^k\otimes\mathbb{K}^n$ and $\mathbb{K}^{k\cdot n}$.

The elements of the dual space to $\mathbb{K}^n$ may be conveniently realised as the space of *row* vectors $\mathsf{w} \in \mathbb{K}^{1\times n}$, i.e. $(\mathbb{K}^n)^* \cong \mathbb{K}^{n\times 1}$, such that the natural duality pairing is $\langle \mathsf{w}, \boldsymbol{x}\rangle := \mathsf{w}\boldsymbol{x} = \boldsymbol{w}^T\boldsymbol{x} \in \mathbb{K}$, where $\boldsymbol{w} := \mathsf{w}^T$. The **dual** basis to the canonical basis $\{\boldsymbol{e}_\jmath\}$ is $\{\mathsf{e}^\imath\}$, where $\mathsf{e}^\imath \in \mathbb{K}^{1\times n}$ is a row vector with a 1 in the $\imath$-th position and 0 elsewhere such that $\langle \mathsf{e}^\imath, \boldsymbol{e}_\jmath\rangle = \mathsf{e}^\imath \boldsymbol{e}_\jmath = \delta^\imath_\jmath$.

The p-norms on $\mathbb{K}^n$ are defined as $\|\boldsymbol{x}\|_p := \left(\sum_{k=1}^n |x_k|^p\right)^{1/p}$ for $1 \le p < \infty$, and by $\|\boldsymbol{x}\|_\infty := \max_k |x_k|$ for $p = \infty$. The associated matrix norms are denoted $\|\boldsymbol{A}\|_p := \max_{\|\boldsymbol{x}\|_p=1} \|\boldsymbol{A}\boldsymbol{x}\|_p$. There is possibility for confusion here, as the *Schatten* p-norms are written similarly (see section 4.6.1), but it should be clear from the context what is meant. The *dual* norm is a special case, and it is the q-norm with $p^{-1} + q^{-1} = 1$, i.e. for $\mathsf{w} \in \mathbb{K}^{1\times n}$ it is $\|\boldsymbol{x}\|_q := \left(\sum_{k=1}^n |x_k|^q\right)^{1/q}$; this extends to $p = 1$ and $q = \infty$ and vice versa.

The **Frobenius-norm** comes from the duality pairing $\langle \boldsymbol{A}, \boldsymbol{B}\rangle := \operatorname{tr} \boldsymbol{B}^*\boldsymbol{A}$, also denoted by $\boldsymbol{A} : \boldsymbol{B}$, such that $\|\boldsymbol{A}\|_F^2 = \operatorname{tr} \boldsymbol{A}^*\boldsymbol{A}$. This is the Schatten 2-norm (see section 4.6.1).

Once a matrix norm is given, **condition numbers** can be defined. The condition number of $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ is $\kappa(\boldsymbol{A}) = \|\boldsymbol{A}\|\|\boldsymbol{A}^{-1}\|$, and it depends on the chosen matrix norm. If this is one of the p-norms, this may be attached to the notation $\kappa_p(\boldsymbol{A}) = \|\boldsymbol{A}\|_p\|\boldsymbol{A}^{-1}\|_p$.

1.9 Matrix Subspaces, Groups and Algebras

The group—w.r.t. matrix multiplication—of invertible matrices in $\mathbb{K}^{n\times n}$ is the general linear group $\mathrm{GL}(n, \mathbb{K})$, also denoted by $\mathrm{GL}(\mathbb{K}^n)$ or just $\mathrm{GL}(n)$. The sub-group with $\det \boldsymbol{A} = 1$ is the special linear group of **unimodular matrices** $\mathrm{SL}(n)$. The group of **orthogonal matrices** is $\mathrm{O}(n)$, and the sub-group of **special orthogonal** matrices is $\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R})$. Similarly for **unitary matrices** $\mathrm{U}(n, \mathbb{C})$, the **special unitary** sub-group is $\mathrm{SU}(n) = \mathrm{U}(n) \cap \mathrm{SL}(n, \mathbb{C})$. Note that the determinant may be seen as a *group homomorphism* $\det : \mathrm{GL}(n, \mathbb{K}) \to \mathbb{K} \setminus \{0\}$ into the multiplicative group of the underlying field, with *kernel* the unimodular matrices.

These are all **Lie groups** w.r.t. matrix multiplication, i.e. topological groups (see section 4.1) which are also differentiable manifolds (see section 11). The corresponding **Lie algebras**, i.e. the tangent space at the neutral element $\boldsymbol{I}$ (see section 11) are denoted by lower case letters, i.e. $\mathrm{gl}(n, \mathbb{K})$ (or, $\mathrm{gl}(\mathbb{K}^n)$ or $\mathrm{gl}(n)$) for the Lie group $\mathrm{GL}(n, \mathbb{K})$. This is the algebra of all matrices $\mathrm{gl}(n, \mathbb{K}) = \mathbb{K}^{n\times n}$ with normal matrix addition and scalar multiplication as vector space, and with the **Lie product** being the **commutator**

$$\boldsymbol{A}, \boldsymbol{B} \mapsto [\boldsymbol{A}, \boldsymbol{B}] := \boldsymbol{A}\boldsymbol{B} - \boldsymbol{B}\boldsymbol{A}.$$

The subspace of **symmetric matrices** ($\boldsymbol{A} = \boldsymbol{A}^T$) in $\mathbb{K}^{n\times n}$, resp. **Hermitian** in the complex case ($\boldsymbol{A} = \boldsymbol{A}^*$), is denoted by $\mathrm{sym}(n)$. The subspace of **skew** ($\boldsymbol{A} = -\boldsymbol{A}^T$) resp. skew Hermitian matrices is denoted by $\mathrm{so}(n)$ resp. $\mathrm{u}(n)$. Matrices $\boldsymbol{A} \in \mathbb{K}^{n\times n}$ where $\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}\boldsymbol{A}^*$ are called **normal.**

One may adopt the recommendation of *Parlett* and use upper case *symmetric letters*—like $\boldsymbol{A}, \boldsymbol{Y}, \boldsymbol{U}$ etc.—for elements of $\mathrm{sym}(n)$ and upper case *unsymmetric letters* for other matrices—like $\boldsymbol{S}, \boldsymbol{E}, \boldsymbol{R}$ etc. Orthogonal matrices like to be called $\boldsymbol{Q}, \boldsymbol{U}, \boldsymbol{V}$ etc., similarly for unitary matrices.

To the Lie group $\mathrm{SL}(n, \mathbb{K})$ corresponds the algebra of trace-less matrices $\mathrm{sl}(n, \mathbb{K}) = \{\boldsymbol{A} \in \mathbb{K}^{n\times n} :\ \mathrm{tr}\ \boldsymbol{A} = 0\}$.

The Lie algebra of $\mathrm{O}(n)$ and $\mathrm{SO}(n)$ is the space of real skew matrices $\mathrm{so}(n) = \{\boldsymbol{A} \in \mathbb{R}^{(n\times n)} :\ \boldsymbol{A} = -\boldsymbol{A}^T\}$.

The Lie algebra of $\mathrm{U}(n)$ is the space of skew Hermitian matrices $\mathrm{u}(n) = \{\boldsymbol{A} \in \mathbb{C}^{(n\times n)} :\ \boldsymbol{A} = -\boldsymbol{A}^*\}$.

The Lie algebra of $\mathrm{SU}(n)$ is the space of traceless skew Hermitian matrices $\mathrm{su}(n) = \{\boldsymbol{A} \in \mathbb{C}^{(n\times n)} :\ \boldsymbol{A} = -\boldsymbol{A}^* \ \wedge\ \mathrm{tr}\ \boldsymbol{A} = 0\}$.

In all cases, if $\boldsymbol{A}$ is from a Lie algebra, $\exp(\boldsymbol{A})$ is in the corresponding Lie group. Note that $\exp(\mathrm{tr}\,\boldsymbol{A}) = \det\exp(\boldsymbol{A})$.

A special case is the space $\mathrm{sym}(n)$: this is a sub-space of $\mathrm{gl}(n)$, and trivially a Lie algebra with the trivial Lie product $[\boldsymbol{A}, \boldsymbol{B}] := \boldsymbol{0}$. The image under $\exp : \mathrm{gl}(n) \to \mathrm{GL}(n)$ is $\exp(\mathrm{sym}(n)) = \mathrm{Sym}^+(n)$, the open cone—and hence sub-manfold of $\mathrm{GL}(n)$—of symmetric resp. Hermitian positive definite (SPD) matrices—$\boldsymbol{x}^*\boldsymbol{A}\boldsymbol{x} > 0$ for $\boldsymbol{x} \neq \boldsymbol{0}$.

These are not a group under normal matrix multiplication, but can be made into a Lie group by noting that $\exp : \mathrm{sym}(n) \to \mathrm{Sym}^+(n)$ is bijective, the inverse being the matrix logarithm: for $\boldsymbol{A} \in \mathrm{Sym}^+(n)$ one has $\exp(\log(\boldsymbol{A})) = \boldsymbol{A}$. One may define a *commutative* group operation $\boxplus$ on $\mathrm{Sym}^+(n)$ by

$$\boldsymbol{A} \boxplus \boldsymbol{H} := \exp(\log(\boldsymbol{A}) + \log(\boldsymbol{H})).$$

The identity is $\boldsymbol{I}$ as $\log(\boldsymbol{I}) = \boldsymbol{0}$, and the inverse of $\boldsymbol{H}$ is $\boxminus\boldsymbol{H} := \exp(-\log(\boldsymbol{H}))$, this makes $\mathrm{Sym}^+(n)$ into an Abelian Lie group with trivial Lie algebra $\mathrm{sym}(n)$. The algebraic structure on $\mathrm{Sym}^+(n)$ can be extended to a linear real vector space structure—mapping it from the Lie algebra via exp—by defining for $\alpha \in \mathbb{R}$ a new scalar multiplication $\boxtimes$ on $\mathbb{R} \times \mathrm{Sym}^+$ by

$$\alpha \boxtimes \boldsymbol{A} := \exp(\alpha \log(\boldsymbol{A})) = \boldsymbol{A}^\alpha,$$

this operation is distributive and continuous, and $0 \boxtimes \boldsymbol{A} = \boldsymbol{A}^0 = \boldsymbol{I}$ and $1 \boxtimes \boldsymbol{A} = \boldsymbol{A}^1 = \boldsymbol{A}$, making $\mathrm{Sym}^+(n)$ into a topolgical vector space isomorphic to $\mathrm{sym}(n) \cong \mathbb{R}^{n\cdot(n+1)/2}$.

1.10 Analysis

A multi-index $\mathfrak{k} = (k_1, \dots, k_m, \dots)$ is an element from either $\mathbb{N}_0^n$ or $\mathcal{N}_{00}$. Then the length $|\mathfrak{k}| := \sum_m k_m$ and the factorial $\mathfrak{k}! := \prod_m (k_m!)$ are well-defined. Special multi-indices are $\{\mathfrak{e}_\jmath\}$, where $\mathfrak{e}_\jmath = (\delta_{\jmath k})_k$ (a 1 in the $\jmath$-th position, 0 otherwise).

Let $\mathcal{K}$ be one of $\mathbb{K}^n$, $\boldsymbol{c}_{00}$, or $\boldsymbol{\omega}$. For $\boldsymbol{x} \in \mathcal{K}$ and a multi-index $\mathfrak{k}$, define the monomials $\boldsymbol{x}^{\mathfrak{k}} := [(x_m^{k_m})_m]^T \in \mathcal{K}$, and corresponding polynomials or power series $\sum_{\mathfrak{k}} a_{\mathfrak{k}} \boldsymbol{x}^{\mathfrak{k}} \in \mathcal{K}$.

The **convolution** of $\boldsymbol{x}, \boldsymbol{y} \in \mathcal{K}$ (with the indices $\mod n$ in case $\mathcal{K} = \mathbb{K}^n$) is defined component-wise as $(\boldsymbol{x} * \boldsymbol{y})_\ell := \sum_m x_m y_{\ell - m}$, and the **correlation** as $(\boldsymbol{x} \star \boldsymbol{y})_\ell := \sum_m x_m y_{\ell + m}$. Note the difference in notation between the *asterisk* "$*$" and the *star* "$\star$". For two functions $f, g \in L_1(\mathbb{K}^n)$ the **convolution** is analogously defined by $(f * g)(\boldsymbol{x}) := \int_{\mathbb{K}^n} f(\boldsymbol{y}) g(\boldsymbol{x} - \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}$, and the **correlation** as $(f \star g)(\boldsymbol{x}) := \int_{\mathbb{K}^n} f(\boldsymbol{y}) g(\boldsymbol{x} + \boldsymbol{y}) \, \mathrm{d}\boldsymbol{y}$.

The derivative for a function $f : \mathbb{K} \to \mathbb{K}$ is denoted by

$$\frac{\mathrm{d}}{\mathrm{d}x} f(x) = \frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{\mathrm{d}f(x)}{\mathrm{d}x} = \mathrm{D}f(x).$$

Observe that the sign for the differential "d" is *not* a variable d, and is denoted differently, as is the derivative sign "D".

For a function $f : \mathcal{K} \to \mathbb{K}$, the partial derivatives w.r.t. x_i have many denotations

$$\frac{\partial f}{\partial x_i} = \partial_{x_i} f = \mathrm{D}_{x_i} f = \mathrm{D}^{\mathfrak{e}_i} f = \partial_i f = \mathrm{D}_i f = f_{,x_i} = f_{,i} \,.$$

For a multi-index $\mathfrak{k}$ the possibly mixed partial derivatives are denoted by

$$\mathrm{D}^{\mathfrak{k}} f(\boldsymbol{x}) := \frac{\partial^{|\mathfrak{k}|}}{\partial x_1^{k_1} \cdots \partial x_m^{k_m} \cdots} f(x_1, \dots, x_m, \dots) := \partial^{\mathfrak{k}} f(\boldsymbol{x}).$$

Note that the derivative $\mathrm{D}f(x) = \mathrm{d}f(x) = (\partial f(x)/\partial x_1, \dots, \partial f(x)/\partial x_k, \dots) \in \mathcal{K}^*$, i.e. if $\mathcal{K} = \mathbb{K}^n$, then $\mathrm{D}f(x) = \mathrm{d}f(x) \in \mathbb{K}^{1 \times n}$ may be represented as a row vector, whereas the gradient $\nabla f(x) := [\mathrm{D}f(x)]^T = [\mathrm{d}f(x)]^T \in \mathbb{K}^n$ is a column vector. Hence $\partial_\jmath f(x) := \mathrm{D}^{\mathfrak{e}_\jmath} f(x) = \mathrm{D}f(x)\boldsymbol{e}_\jmath = \mathrm{d}f(x)\boldsymbol{e}_\jmath = (\nabla f(x))^T \boldsymbol{e}_\jmath$, showing that $\mathrm{D}f(x) = \mathrm{d}f(x) = \sum_\jmath \partial_\jmath f(x) \mathbf{e}^\jmath$, but $\nabla f(x) = \sum_\jmath \partial_\jmath f(x) \boldsymbol{e}_\jmath$. The gradient operator ∇ is also called the **nabla** operator.

The **Hessian** is the symmetric matrix of second derivatives

$$\mathrm{D}^2 f(x) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2}(x) & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(x) \\ \dots & \frac{\partial^2 f}{\partial x_\imath \partial x_\jmath}(x) & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x) \end{bmatrix} = \left[\frac{\partial^2 f}{\partial^{k_1} x_1 \cdots \partial^{k_n} x_n} \right]_{k_1, \dots, k_n = 0}^{|\mathfrak{k}| = 2} .$$

The **Laplace** operator is $\Delta f(x) := \operatorname{tr} \mathrm{D}^2 f(x)$, sometimes also written as the square of the nabla operator $\nabla^2 fx) = \Delta f(x)$.

Taylor's formula reads

$$f(\boldsymbol{x}+\boldsymbol{h}) = \sum_{\mathfrak{k}} \frac{1}{\mathfrak{k}!} \mathrm{D}^{\mathfrak{k}} f(\boldsymbol{x}) \boldsymbol{h}^{\mathfrak{k}} + R(\boldsymbol{x},\boldsymbol{h}).$$

For a function $\boldsymbol{f} : \mathbb{K}^n \to \mathbb{K}^m$, $\boldsymbol{f}(\boldsymbol{x}) = \sum_\imath f^\imath(\boldsymbol{x}) \boldsymbol{e}_\imath$, the derivative or **Jacobi** matrix is

$$\mathrm{D}\boldsymbol{f}(\boldsymbol{x}) = \begin{bmatrix} \frac{\partial f^1}{\partial x_1}(x) & \dots & \frac{\partial f^1}{\partial x_n}(x) \\ \dots & \frac{\partial f^\imath}{\partial x_\jmath}(x) & \dots \\ \frac{\partial f^m}{\partial x_1}(x) & \dots & \frac{\partial f^m}{\partial x_n}(x) \end{bmatrix} = \left[\frac{\partial f^\imath}{\partial x_\jmath} \right]_{\jmath=1,\dots,n}^{\imath=1,\dots,m}.$$

For $m = n$ this is a square matrix and the determinant is called the **Jacobian**: $\det \mathrm{D}\boldsymbol{f}(\boldsymbol{x})$.

Higher order derivatives are tensors of higher degree, e.g. the k-th derivative of a function $\boldsymbol{f} : \mathbb{K}^n \to \mathbb{K}^m$ is a tensor of degree $k+1$; with multi-indices $\mathfrak{k} = (k_1, \dots, k_n) \in \mathbb{N}_0^n$:

$$\mathrm{D}^k \boldsymbol{f}(\boldsymbol{x}) = \left[\frac{\partial^k f^\imath}{\partial x_{k_\jmath}} \right]_{|\mathfrak{k}|=k}^{\imath=1,\dots,m}.$$

2 Sets with Convergence Structures

This will include general topological, uniform, pseudo-metric, and metric spaces. The combination with measures will come in the section 3, and the combination of topological with algebraic structures will come in section 4.1.

2.1 Topology

For collections of sets one may use capital letters in *Fraktur* font, e.g. if on a set S a topology is defined, this may be denoted by $\mathfrak{T} \subseteq 2^S$. Each topology $\mathfrak{T}$ contains both S and $\emptyset$, and is closed under *finite intersections* and *arbitrary unions.* The sets in $\mathfrak{T}$ are the open sets O, where the interior $\mathring{O} := \operatorname{int} O = O$. The complements of open sets are closed sets C, where the closure $\bar{C} := \operatorname{cl} C = C$. The **boundary** of a set $A \subset S$ is $\partial A := \operatorname{cl} A \setminus \operatorname{int} A$, a closed set. If $s \in S$, the **neighbourhoods** of s —the neighbourhood filter, see section 1.2—are $\mathfrak{U}(s) = \{V \subseteq S |\ \exists O \in \mathfrak{T} :\ s \in O \subseteq V\}$. If these filters have a *countable* basis (see section 1.2) for all $s \in S$, the topological space $(S, \mathfrak{T})$ is called **first countable**.

A point $s \in S$ is a **closure point** of a subset $A \subseteq S$, if $A \cap U \neq \emptyset$ for each $U \in \mathfrak{U}(s)$. The set of all closure point is $\operatorname{cl} A$. This is the smallest closed set which contains A, or equivalently the intersection of all such sets. The point s is an **accumulation**, **limit**, or **cluster point** of A if it is a closure point of $A \setminus \{s\}$. The set of all accumulation points is the **derived** set, denoted by A'. The point s is an **interior point** if there is a $U \in \mathfrak{U}(s)$ such that $U \subseteq A$. The set of all interior points is $\operatorname{int} A$. This is the largest open set contained in A, or equivalently the union of all such sets. The **boundary points** of A are those points which are closure points of both A and $\complement A$. The set of all boundary points is ∂A. Hence $\partial A = \partial \complement A = \operatorname{cl}(A) \cap \operatorname{cl}(\complement A) = \operatorname{cl}(A) \setminus \operatorname{int}(A)$. A set $B \subseteq A$ is called **dense** in A if $\operatorname{cl} B = A$. In case there is a *countable* set which is *dense* in S, the topological space $(S, \mathfrak{T})$ is called **separable**.

A family of sets $\mathfrak{Q} \subseteq 2^S$ is called a **basis** for the topology, if any open set is the union of sets in $\mathfrak{Q}$, i.e.

$$O \in \mathfrak{T} \;\Rightarrow\; O = \bigcup_{Q \in \mathfrak{Q}'} Q, \; \mathfrak{Q}' \subseteq \mathfrak{Q}.$$

The topology generated by $\mathfrak{Q}$ in this way (by arbitrary unions) is denoted by $\mathfrak{T}(\mathfrak{Q})$. Note that $\mathfrak{Q}$ can only be a basis of a topology if it is a covering of S, and if for all $Q_1, Q_2 \in \mathfrak{Q}$ and any $s \in Q_1 \cap Q_2$ there is a $R \in \mathfrak{Q}$ with $s \in R \subseteq Q_1 \cap Q_2$. A covering $\mathfrak{P}$ of S is a **sub-basis** of the topology $\mathfrak{T}$ if the finite intersections of sets in $\mathfrak{P}$ are a *basis*. The generated topology is again denoted by $\mathfrak{T}(\mathfrak{P})$. Note that any covering is a *sub-basis* for a topology. If the topological space $(S, \mathfrak{T})$ has a *countable sub-basis*, it is called **second countable**.

A subset $B \subseteq S$ is called **compact**, if every open cover $\bigcup_j A_j \supseteq B$ contains a finite sub-cover. A set is called **relatively compact** or **pre-compact** if its closure is compact. Note that in a compact set every filter or net has a limit point. The space S is called **locally compact**, if every point $s \in S$ has a compact neighbourhood. A collection of subsets $\{A_j\}_{j \in J}$ of S is called **locally finite** if each point $s \in S$ has an open neighbourhood which meets only finitely many of the A_j. A set B is called **para-compact** if every open cover has an open locally finite refinement cover. Observe that every metrisable space is para-compact. A space S is called σ-**compact** if it is the countable union of compact sets. A set B is called a $\mathcal{G}_\delta$-set if it is the intersection of countably many open sets. A set C is a $\mathcal{F}_\sigma$-set if it is the union of countably many closed sets.

If for any two points $r \neq s \in S$ there are neighbourhoods $U_r \in \mathfrak{U}(r)$ and $U_s \in \mathfrak{U}(s)$ such that $U_r \cap U_s = \emptyset$, then the topological space S is called a **Hausdorff** space or T_2-space.

Let J be a directed set (see section 1.5), then an element of S^J is called a **net**; usually denoted by $\{x_j\}_{j\in J}$. In case $J = \mathbb{N}$, the net is called a **sequence.** The net **converges** to a **limit** $s \in S$ if for each $U \in \mathfrak{U}(s)$ there is a $j_U \in J$ such that $x_j \in U$ if $j_U \preceq j$, we write $s = \lim_{j\in J} s_j$. If S is Hausdorff, the limit is unique. Related to a net (or sequence) is the *filter basis* of sets of the form $\mathcal{F}_{j_0} = \{x_j : \; j_0 \preceq j\}$ for all $j_0 \in J$, generating the **tail filter**. A filter (not just the tail filter) converges to a point $s \in S$, if it contains the neighbourhood filter $\mathfrak{U}(s)$. Hence the neighbourhood filter $\mathfrak{U}(s)$ is the smallest filter which converges to s. A point $s \in S$ is a **limit** or **accumulation** point of a filter $\mathfrak{F}$ if $s \in \operatorname{cl} A$ for each $A \in \mathfrak{F}$, denoted by $s \in \lim \mathfrak{F}$.

A map $\phi \in \mathscr{F}(\mathcal{X}, \mathcal{Y})$ between the two topological spaces $\mathcal{X}$ and $\mathcal{Y}$ is **continuous at** $x \in \mathcal{X}$ if for each $U \in \mathfrak{U}(\phi(x))$ there is a $V \in \mathfrak{U}(x)$ such that $\phi(V) \subseteq U$. A map is everywhere continuous or simply **continuous** if it is so for all $x \in \mathcal{X}$. The everywhere continuous maps—that is maps $\phi \in \mathscr{F}(\mathcal{X}, \mathcal{Y})$ such that $\phi^{-1}(O) \in \mathfrak{T}_{\mathcal{X}}$ for each $O \in \mathfrak{T}_{\mathcal{Y}}$—are denoted by $C(\mathcal{X}, \mathcal{Y})$, if $\mathcal{Y} = \mathbb{K}$ this is usually written as $C(\mathcal{X})$. Special subsets are $C_{00}(\mathcal{X}) \subseteq C(\mathcal{X})$, (also denoted as $C_c(\mathcal{X})$ and (**beware**) also as $C_0(\mathcal{X})$) the space of all continuous functions f with compact support

$$\operatorname{supp} f = \operatorname{cl}\{x \in \mathcal{X} |\; f(x) \neq 0\},$$

and the space of continuous functions which "vanish at infinity" $C_0(\mathcal{X}) \subseteq C(\mathcal{X})$, where $\{x \in \mathcal{X} : \; |f(x)| \geq \varepsilon\}$ is compact $\forall \varepsilon > 0$, and the set of bounded continuous functions $C_b(\mathcal{X}) := \{f \in C(\mathcal{X}) : \; \sup_{x\in\mathcal{X}} |f(x)| < \infty\}$. Obviously $C_{00}(\mathcal{X}) \subseteq C_0(\mathcal{X}) \subseteq C_b(\mathcal{X}) \subseteq C(\mathcal{X})$. Another space which is sometimes important are the bounded—not necessarily continuous—functions $\mathscr{B}(\mathcal{X}) := \{f \in \mathbb{R}^{\mathcal{X}} : \; \sup_{x\in\mathcal{X}} |f(x)| < \infty\}$, so that $C_b(\mathcal{X}) = \mathscr{B}(\mathcal{X}) \cap C(\mathcal{X})$. With the co-domain $\mathcal{Y} = \mathbb{K}$, note that all these spaces are are subsets of $\mathscr{F}(\mathcal{X}, \mathbb{K})$ and hence are *algebras* under pointwise operations.

A map $\phi \in \mathscr{F}(\mathcal{X}, \mathcal{Y})$ is called **open**, iff $\phi(O) \in \mathfrak{T}_{\mathcal{Y}}$ for each $O \in \mathfrak{T}_{\mathcal{X}}$. In case $\phi \in C(\mathcal{X}, \mathcal{Y})$ is *bijective* and *open*, it has a continuous inverse $\phi^{-1} \in C(\mathcal{Y}, \mathcal{X})$, and is called a **homeomorphism**.

Let $(S, \mathfrak{T}_S)$ be a topological space, A any set, and $\phi \in \mathscr{F}(A, S)$. Then the family of sets $\mathfrak{T}_A = \{\phi^{-1}(O) |\; O \in \mathfrak{T}_S\}$ is the *weakest* or *coarsest* topology on A such that ϕ is continuous. It is called the **initial** topology, or the topology **induced** by ϕ. If $Q \subset S$, then the *inclusion* $\iota : Q \hookrightarrow S$ induces a topology on Q, called the **relative topology**. Hence a subset $R \subseteq Q$ is relatively open iff there is an open set $O \in \mathfrak{T}_S$ such that $R = Q \cap O$.

If $\{S_j\}_{j\in J}$ is a family of topological spaces, then the product (see section 1.2) $\prod_{j\in J} S_j$ may be equipped with the *initial* topology such that all the canonical projections are continuous, this is the **product topology.** The canonical projections are also *open.*

Conversely, let $\psi \in \mathscr{F}(S,A)$. Then the family of sets $\mathfrak{T}_A = \{H \mid H \subseteq A,\ \psi^{-1}(H) \in \mathfrak{T}_S\}$ is the *strongest* or *finest* topology on A such that ψ is continuous. It is called the **final** topology, or the **identification** topology of ψ. If $r \sim s$ is an equivalence relation (see section 1.4) on S, and $q : S \to S/\sim$ the natural projection onto the quotient, then $S/\sim$ may be equipped with the *final* or *identification* topology such that q is continuous. The equivalence relation is called open if the natural projection is open.

If $\{S_j\}_{j\in J}$ is a family of topological spaces, then the disjoint union $\biguplus_{j\in J} S_j$ (see section 1.2) may be equipped with the *final* topology such that all the canonical injections are continuous, this is the **sum topology** and the disjoint union is called the **topological sum**, sometimes denoted by $\sum_{j\in J} S_j$, or also $\coprod_{j\in J} S_j$. The canonical injections are also *open.*

As $\mathscr{F}(A,\mathcal{Y})$ may be written also as $\mathcal{Y}^A$, where $\mathcal{Y}$ is a topological space and A any set, it may be equipped with the *product topology.* This is here also called the topology of **pointwise convergence** or the **simple topology**. This topology in turn is *induced* on all subspaces of $\mathscr{F}(A,\mathcal{Y})$. Especially $\mathscr{B}(\mathcal{X})$ and its subspaces $C(\mathcal{X}), C_b(\mathcal{X}), C_0(\mathcal{X})$, and $C_{00}(\mathcal{X})$ may be equipped with this topology, but they are not closed subspaces with this topology. The closure of $C(\mathcal{X})$ in this topology in $\mathscr{F}(\mathcal{X},\mathbb{R})$ is called the class of **Baire functions**, i.e. pointwise limits of continuous functions. More precisely, the continuous functions are called the *Baire class*-0 functions, and their limits are the *Baire class*-1. This process can be continued inductively to *Baire class-n* for any $n \in \mathbb{N}$; this is a growing *nested* sequence. A function φ is said to be of *Baire class-m*, if m is the smallest integer for which φ is in the *Baire class-m*. Note that there is functions of every *Baire class-m* for all $m \in \mathbb{N}$. The set of all these functions is then the **Baire class**.

On $C(\mathcal{X},\mathcal{Y})$ one may define the **compact open** topology with a basis of sets $\{\phi \in C(\mathcal{X},\mathcal{Y}) :\ \phi(\mathcal{K}) \subset \mathcal{U},\ \mathcal{K} \subseteq \mathcal{X} \text{ compact},\ \mathcal{U} \subseteq \mathcal{X} \text{ open}\}$, it is the uniform convergence on compact sets. $C(\mathcal{X})$ and $C_{00}(\mathcal{X})$ are complete spaces with this topology. $C_b(\mathcal{X})$ and $C_0(\mathcal{X})$ are Banach spaces with the norm $\|\phi\|_\infty = \sup_{x\in\mathcal{X}} |\phi(x)|$.

2.2 Metric and Pseudo-Metric Spaces

Often the topology may be generated by a **pseudo-metric**, i.e. a symmetric function $d_S : S \times S \to \mathbb{R}_+$ with $d_S(s,s) = 0$ for all $s \in S$ which satisfies the **triangle inequality** $\forall r,s,t \in S :\ d_S(r,s) \le d_S(r,t) + d_S(t,s)$, where $d_S(r,s)$ is called the *distance* between r and s.

By $B_{d_S}(s,\epsilon) := \{r \in S : \; d_S(s,r) < \epsilon\}$ denote the (formally) open **ball** of radius ϵ around $s \in S$, and by $\bar{B}_{d_S}(s,\epsilon) := \{r \in S : \; d_S(s,r) \leq \epsilon\}$ the corresponding (formally) closed ball. If the metric is clear from the context, this is also denoted by just $B(s,\epsilon)$, etc. The sets of all balls $\mathfrak{N}(s) := \{B_{d_S}(s,\epsilon) : \epsilon > 0\}$ around each $s \in S$ form a *filterbasis* for the neighbourhood filter $\mathfrak{U}(s)$. A set $R \subseteq S$ is then open if for each point $s \in R$ there is an $\epsilon > 0$ such that $B_{d_S}(s,\epsilon) \subseteq R$. A topological space whose topology may be generated by a pseudo-metric is called *pseudo-metrisable*.

A sequence $\{x_j\}_{j\in\mathbb{N}}$ in a pseudo-metric space converges to its limit $s \in S$ iff for any $B(s,\epsilon)$ there is a $n \in \mathbb{N}$ such that for all $m > n$ one has $x_m \in B(s,\epsilon)$; equivalently $d_S(x_m, s) < \epsilon$, one writes $s = \lim_{j\to\infty} x_j$ or $x_j \to s$.

In case $r \neq s \Rightarrow d_S(r,s) > 0$, the pseudo-metric is called a **metric**. The best known example is $\mathbb{R}$ with $d(r,s) := |r - s|$. The resulting topological space is then a *Hausdorff* space, hence a limit of a sequence in a metric space is unique. A topological space whose topology may be generated by a metric is called *metrisable*.

A collection $\mathcal{D} = \{d_j\}_{j\in J}$ of pseudo-metrics is called a **gauge system**. One says that the gauge system **separates** the points of S, if $r \neq s \Rightarrow \exists j \in J : d_j(r,s) > 0$. Note that if d_S is a metric, then the gauge system $\{d_S\}$ separates the points. Let $\mathfrak{T}_\mathcal{D} \subseteq 2^S$ be such that for each $R \in \mathfrak{T}_\mathcal{D}$ and all $s \in R$ there is a finite $J_0 \subseteq J$ and an $\epsilon > 0$ such that $\bigcap_{j\in J_0} B_{d_j}(s,\epsilon) \subseteq R$. Then $\mathfrak{T}_\mathcal{D}$ is the topology generated by the gauge system. If the gauge system separates the points, the topology is Hausdorff. In case the set J is countable—hence may be replaced by $\mathbb{N}$ —and the gauge system separates the points, the topology and uniformity of the space may be generated by a metric, e.g.

$$d(r,s) := \sum_{n\in\mathbb{N}} \frac{\min\{d_n(r,s), 1\}}{2^n}.$$

The metric space S is called **bounded** if for some $\rho \in \mathbb{R}_*$ one has $d(r,s) \leq \rho$ for all $r,s \in S$. The smallest possible such ρ is called the **diameter** of S. If in addition for any $\rho > 0$ there are finitely many $s_j, j = 1,\ldots,m$ such that the balls $B(s_j,\rho)$ cover S, the space is called **pre-compact** or **totally bounded**. Every totally bounded space is bounded.

A concept that is generalised in section 2.3 is that of a **Cauchy sequence** $\{x_j\}_{j\in\mathbb{N}}$, where for any $\epsilon > 0$ there is a $n \in \mathbb{N}$, such that for any $k, m > n$ one has $d_S(x_k, x_m) < \epsilon$. Obviously every convergent sequence is a Cauchy sequence. In case that every Cauchy sequence converges, the metric space S is called **complete**. One may embed any metric space densely in a complete metric space, which is then called its **completion**.

One may define a distance between a point $s \in S$ and a subset $Q \subseteq S$ by

$d(s, Q) := \inf\{d(s, r) : r \in Q\}$. Then $d(s, Q) = 0$ iff $s \in \operatorname{cl} Q$. As for any $r \in S$ one has $d(s, Q) \leq d(s, r) + d(r, Q)$, the map $s \mapsto d(s, Q)$ is continuous. With this concept one can define the **Hausdorff distance** between two subsets $Q, R \subseteq S$:

$$d_H(Q, R) := \max\{\sup_{s \in Q} d(s, R), \sup_{s \in R} d(a, Q)\}.$$

The Hausdorff distance d_H makes the set $\mathfrak{K}(S)$ of all non-empty compact subsets of S into a metric space. $\mathfrak{K}(S)$ is complete if S is complete.

If (S, d_S) and (T, d_T) are metric spaces, a map ϕ is **continuous** at $s \in S$ iff for a convergent sequence $x_j \to s$ one has that $\phi(x_j) \to \phi(s)$. If this holds at any $s \in S$, the map is continuous, $\phi \in C(S, T)$. A map is **uniformly continuous** (see also section 2.3) at $s \in S$ iff for any $\epsilon > 0$ there is a $\delta > 0$ such that $d_S(r, s) < \delta \Rightarrow d_T(\phi(r), \phi(s)) < \epsilon$. A Cauchy sequence with limit $s \in S$ is mapped into a Cauchy sequence with limit $\phi(s) \in T$ by a uniformly continuous map. If this holds at any $s \in S$, the map is uniformly continuous, denoted by $C_u(S, T)$. A uniformly continuous map is obviously continuous.

A subset $\mathscr{H} \subset C(S, T)$ is called **equicontinuous** at $s \in S$ iff for any $\epsilon > 0$ there is a $\delta > 0$ such that for all $\phi \in \mathscr{H}$ one has $d_S(r, s) < \delta \Rightarrow d_T(\phi(r), \phi(s)) < \epsilon$. If this holds at all $s \in S$, the set $\mathscr{H}$ is called just equicontinuous. Obviously then all the maps $\phi \in \mathscr{H}$ are uniformly continuous. If one can pick δ independently of $s \in S$, the set $\mathscr{H}$ is called **uniformly equicontinuous**.

A uniformly continuous map $\phi \in C_u(S, T)$ is an **isometry** iff $\forall r, s \in S :$ $d_S(r, s) = d_T(\phi(r), \phi(s))$. In case $d_T(\phi(r), \phi(s)) \leq L\, d_S(r, s)$ for all $r, s \in S$, the map ϕ is called **Lipschitz continuous** with **Lipschitz constant** $L \geq 0$; it is obviously continuous, and the space of Lipschitz continuous maps is denoted by $C^{0,1}(S, T) \subset C_u(S, T) \subset C(S, T)$. The **modulus of continuity** of $\phi \in \mathscr{F}(S, T)$ is

$$\omega_\phi(\epsilon) := \sup\{d_T(\phi(r), \phi(s)) :\ (r, s) \in S^2,\ d_S(r, s) \leq \epsilon\}.$$

For *Lipschitz continuous* ϕ one has $\omega_\phi(\epsilon) \leq L\, \epsilon$, where $L \geq 0$ is the Lipschitz constant of $\phi \in C^{0,1}(S, T)$. Isometries allow $L = 1$, and if $L < 1$ the map is called a **contraction**.

More generally, for such a map ϕ and $0 \leq \beta \leq 1$, let

$$|\phi|_\beta := \sup\left\{ \frac{d_T(\phi(r), \phi(s))}{d_S(r, s)^\beta} \;\middle|\; r \neq s \in S \right\}.$$

One may then define the class of *Hölder* β-continuous functions by

$$C^{0,\beta}(S, T) = \{\phi \in C(S, T) : \quad |\phi|_\beta < \infty\}.$$

Another way of stating $\phi \in C^{0,\beta}(S,T)$ is to say $\omega_\phi(\epsilon) \leq L\,\epsilon^\beta$. Obviously, for $\beta = 1$ the maps are Lipschitz continuous, and for $\beta = 0$ they may be called bounded, i.e. $C_b(S) = C^{0,0}(S,\mathbb{R})$.

One may define a metric on the set $\mathscr{F}(S,T)$ of all maps from any set S into a metric space T by

$$d_\Phi(\phi,\psi) := \sup\{\min\{d_T(\phi(s),\psi(s)),1\}| \quad s \in S\}.$$

The topology so induced is called the topology of **uniform convergence**. If T is complete, so is the space of functions with this metric. If S is a topological space or a metric space, this is a metric on the closed subspaces $C(S,T)$ or $C_u(S,T)$. Again if T is complete, so are these spaces. If a sequence $\{\phi_j\} \subset C_u(S,T)$ is equicontinuous and converges pointwise (at each $s \in S$), it converges uniformly and the limiting function is hence continuous. More generally, *Ascoli's* theorem says that a subset $\mathscr{H} \subset C(S,T)$ is pre-compact iff it is pointwise bounded and equicontiuous.

Hence on $\mathscr{F}(S,\mathbb{R})$ and its subspaces $\mathscr{B}(S), C(S), C_b(S), C_0(S), C_u(S)$, and $C_{00}(S)$, d_Φ is a possible metric, defining the notion of uniform convergence. Observe that all the just mentioned spaces except the last are closed and hence complete subspaces of $\mathscr{F}(S,\mathbb{R})$ with this topology. $C_{00}(S)$ is dense in $C_0(S)$ in this topology.

2.3 Uniformity

Let $\mathfrak{W} \subseteq 2^{S\times S}$ be a *proper filter* such that each $\mathcal{U} \in \mathfrak{W}$, $\mathcal{U} \subseteq S \times S$ is a *reflexive relation* on the set S, for each $\mathcal{U} \in \mathfrak{W}$ there is a $\mathcal{V} \in \mathfrak{W}$ with $\mathcal{V} \subseteq \mathcal{U}^{-1}$, and there is a $\mathcal{W} \in \mathfrak{W}$ with $\mathcal{W} \circ \mathcal{W} \subseteq \mathcal{U}$. Then $\mathfrak{W}$ is called a **uniformity** on S, and the elements $\mathcal{U} \in \mathfrak{W}$ are called **entourages.**

In case d_S is a pseudo-metric on S, define for $\epsilon > 0$ the sets $\mathcal{U}_\epsilon = \{(r,s) \in X \times X : \ d_S(r,s) < \epsilon\}$. Then the collection of these sets is a *filterbasis* of a uniformity, generated by the pseudo-metric. Likewise, if $\mathcal{D} = \{d_j\}_{j\in J}$ is a gauge system, define the sets $\mathcal{U}_{\epsilon,j}$ analogously to $\mathcal{U}_\epsilon$. Then the set of finite intersections of such $\mathcal{U}_{\epsilon,j}$ is a *filterbasis* of a uniformity, generated by the gauge system $\mathcal{D}$. Note that any uniformity may be generated by a gauge system.

For a uniformity $\mathfrak{W}$, define for each $r \in S$ the sets $U_{r,\mathcal{U}} = \{s \in S : \ (r,s) \in \mathcal{U} \in \mathfrak{W}\}$. Then the collection $\{U_{r,\mathcal{U}}\}_{\mathcal{U}\in\mathfrak{W}}$ of all such sets is a *filterbasis* for the *neighbourhood filter* $\mathfrak{U}(r)$ of $r \in S$. The topology so generated is called a **uniform topology**. A topology which may be generated in this way is called *uniformisable.* Each map $r \mapsto U_{r,\mathcal{U}}$ is called a **covering system**—as the $U_{r,\mathcal{U}}$ cover S —and a collection of covering systems—here by varying the

$\mathcal{U} \in \mathfrak{W}$ —is a **neighbourhood system**. In turn each such neighbourhood system may determine a uniformity through its covering systems. In case the uniformity comes from a metric d, the essential covering systems look like $r \mapsto B_d(r, \epsilon)$.

A continuous map $\phi \in C(S,T)$ between uniform topological spaces S and T is called **uniformly continuous** if $(\phi \times \phi)^{-1}(\mathcal{U}) \in \mathfrak{W}_S$ for each $\mathcal{U} \in \mathfrak{W}_T$. Alternatively, for any given covering system $\mathcal{P}$ of the uniformity $\mathfrak{W}_T$ there is a covering system $\mathcal{C}$ of the uniformity $\mathfrak{W}_S$ such that $\forall s \in S$: $\phi(\mathcal{C}(s)) \subseteq \mathcal{P}(\phi(s))$. The space of uniformly continuous functions is denoted by $C_u(S,T) \subseteq C(S,T)$. In case the uniformity is generated by a metric, the condition for uniform continuity may be written in the familiar way as $\forall \epsilon > 0 \, \exists \delta > 0 : \; d(r,s) < \delta \Rightarrow d(\phi(r), \phi(s)) < \epsilon$. Observe that a *Lipschitz continuous* map is *uniformly continuous*, $C^{0,1}(S,T) \subset C_u(S,T)$.

A subset $\mathscr{H}$ of $C(S,T)$ is called **equicontinuous** at $s \in S$ in case for any given covering system $\mathcal{P}$ of the uniformity $\mathfrak{W}_T$ there is a covering system $\mathcal{C}$ of the uniformity $\mathfrak{W}_S$ such that $\forall \phi \in \mathscr{H}$ it holds that $\phi(\mathcal{C}(s)) \subseteq \mathcal{P}(\phi(s))$. Again, in case the uniformity is generated by a metric, this is equivalent to $\forall \epsilon > 0 \, \exists \delta > 0 \, \forall \phi \in \mathscr{H} : \; d(r,s) < \delta \Rightarrow d(\phi(r), \phi(s)) < \epsilon$. If $\forall \phi \in \mathscr{H}$, $\forall s \in S$ it holds that $\phi(\mathcal{C}(s)) \subseteq \mathcal{P}(\phi(s))$, the set $\mathscr{H}$ is **uniformly equicontinuous.** In a metric space this reads as $\forall \epsilon > 0 \, \exists \delta > 0 \, \forall r, s \in S \, \forall \phi \in \mathscr{H} : \; d(r,s) < \delta \Rightarrow d(\phi(r), \phi(s)) < \epsilon$.

A proper filter $\mathfrak{F}$ is a **Cauchy** filter if for every $\mathcal{U} \in \mathfrak{W}$ there is a $F \in \mathfrak{F}$ such that $F \times F \subseteq \mathcal{U}$. Nets and sequences are called Cauchy nets or Cauchy sequences respectively, if the corresponding *tail filter* (see sections 1.2 and 2.1) is a Cauchy filter. Note that every convergent filter is a Cauchy filter. On the other hand, a topological space where every Cauchy filter is convergent is called **complete.** In case the uniformity is generated by a metric d, this means that for every $\epsilon > 0$ there is a j_0, such that for $j_0 \leq j, k$ it holds that $d(x_j, x_k) < \epsilon$.

3 Measure Spaces

A nonempty collection of subsets $\mathfrak{R} \subseteq 2^{\Omega}$ of a set Ω is a **ring** iff it is closed under finite unions and complements, i.e. with $(A, B \in \mathfrak{R}) \Rightarrow (A \cup B \in \mathfrak{R}) \wedge (A \setminus B \in \mathfrak{R})$. It is a ring in the algebraic sense with addition taken as $A \vartriangle B$ and multiplication as $A \cap B$. Obviously the empty set $\emptyset$ is the neutral element for "addition". In case also $\Omega \in \mathfrak{R}$, then the ring is an **algebra**, and clearly Ω is the neutral element for "multiplication". In case the ring is closed under **countable unions**, it is called a **σ-ring**. Again, if $\Omega \in \mathfrak{R}$, the σ-ring is called a **σ-algebra**. The tuple $(\Omega, \mathfrak{R})$ is called a **measurable**

space.

3.1 Charges and Measures

If $\mathfrak{R}$ is an algebra, a function $\mu : \mathfrak{R} \to \overline{\mathbb{R}}$ which is additive—i.e. if $\bigcup_j A_j = A$ is a finite partition of A from $\mathfrak{R}$ then $\mu(A) = \mu(\bigcup_j A_j) = \sum_j \mu(A_j)$ —and assumes at most one of the values $\pm\infty$ and satisfies $\mu(\emptyset) = 0$ is called a **signed charge**. If $\forall A \in \mathfrak{R}: \ \mu(A) \geq 0$, the function is simply called a (positive) **charge**. Note that the charges are a positive cone in the space of signed charges $\boldsymbol{ba}(\Omega, \mathfrak{R})$.

For a charge μ, if $\mathfrak{A}$ is a σ-algebra, and the charge μ is **σ-additive**—i.e. if $\bigcup_j A_j = A$ is an at most countable partition of A from $\mathfrak{A}$ then $\mu(A) = \mu(\bigcup_j A_j) = \sum_j \mu(A_j)$ —then it is called a **signed measure**. If $\forall A \in \mathfrak{A}$: $\mu(A) \geq 0$, the function is just called a (positive) **measure**. Note that the measures are a positive cone in the space of signed measures $\boldsymbol{ca}(\Omega, \mathfrak{A})$. By $(\Omega, \mathfrak{A}, \mu)$ a **measure space** is denoted.

A measure space $(\Omega, \mathfrak{A}, \mu)$ is called **finite** if $\mu(\Omega) < \infty$. It is called **σ-finite**, if it may be covered by the union of a countable number of subsets with finite measure each. Finite measure spaces with $\mu(\Omega) = 1$ are called **probability spaces**; the measure space is often denoted by $(\Omega, \mathfrak{A}, \mathbb{P})$, i.e. the probability measure is denoted as $\mathbb{P}$ with $\mathbb{P}(\Omega) = 1$.

Sets $N \in \mathfrak{A}$ with $\mu(N) = 0$ are called **null-sets**. A measure space $(\Omega, \mathfrak{A})$ is called **complete**, if each subset of a null-set is measurable and is also a null-set. Every measure space may be completed by considering the σ-algebra generated by $\mathfrak{A}$ and the *null-sets*, and this is denoted by $\mathfrak{A}_\mu$. A condition in a measurable space is said to hold **almost everywhere** (abbreviated **a.e.**), if it holds in $\Omega \setminus N$, where N is a null-set. In stochastics, this is also termed as being valid **almost surely** (abbreviated **a.s.**). A subset $A \in \mathfrak{A}$ in a complete measure space is an **atom** of the measure μ if $\mu(A) \neq 0$, and for every subset $B \subseteq A$ either $\mu(B) = 0$ or $\mu(A \setminus B) = 0$. If μ has no atoms, it is called non-atomic or atom-less. A measure μ is purely atomic if there is a countable $A \in \mathfrak{A}$ with $\mu(\Omega \setminus A) = 0$, such that for any $a \in A$ one has $\mu(\{a\}) \neq 0$.

3.2 Measurable Functions

A function $f \in \mathscr{F}(\Omega, \mathbb{R})$ is called **measurable**, if $f^{-1}(B) \in \mathfrak{A}$ for every $B \in \mathfrak{B}$ (the Borel algebra of $\mathbb{R}$). The measurable functions are denoted by $L_0(\Omega, \mathfrak{A}, \mathbb{R}) \subset \mathscr{F}(\Omega, \mathbb{R})$, or just $L_0(\Omega)$ if the σ-algebra $\mathfrak{A}$ is clear from the context—more precisely equivalence classes of maps equal a.e. Slightly more general, a map $\phi \in \mathscr{F}(\Omega, \Xi)$ —where $(\Omega, \mathfrak{A})$ and $(\Xi, \mathfrak{C})$ are measurable spaces—is called measurable($g \in L_0(\Omega, \Xi) \subseteq \mathscr{F}(\Omega, \Xi)$), if

$\forall C \in \mathfrak{C}: \ \phi^{-1}(C) \in \mathfrak{A}$. The measure $\forall C \in \mathfrak{C}: \ g_*\mu(C) := \mu(g^{-1}(C))$ is called the **push-forward** or **image** measure of μ. In *probability theory* this measure (which for real-valued functions—random variables—is a measure on $\mathbb{R}$) is called the **law of** g.

The σ-algebra generated by a collection $\mathfrak{K} \subseteq 2^{\Omega}$ —the smallest σ-algebra containing $\mathfrak{K}$, or generated by a measurable function f (or a set of functions) via $\mathfrak{K} = \{f^{-1}(B) : \ B \in \mathfrak{B}\}$, is denoted by $\Sigma(\mathfrak{K})$, or $\Sigma(f)$ respectively.

If $\{(\Omega_j, \mathfrak{A}_j, \mathbb{P}_j)\}_{j \in J}$ is a collection of *probability spaces*, then there is a smallest σ-algebra on the product $\prod_{j \in J} \Omega_j$ which makes all canonical projections measurable. This is the **product σ-algebra**. The **product measure** is $\mathbb{P} = \bigotimes_{j \in J} \mathbb{P}_j$, such that for $A \subseteq \prod_{j \in J} \Omega_j$ with $A = \prod_{k \in J} A_k$ one has $\mathbb{P}(A) = \bigotimes_{j \in J} \mathbb{P}_j(\prod_{k \in J} A_k) = \prod_{j \in J} \mathbb{P}_j(A_j)$.

3.3 Decomposition of Measures

For an algebra (in case of charges) or a σ-algebra (in case of measures) $\mathfrak{R}$, a set $A \in \mathfrak{R}$ is called **positive** w.r.t. the charge or measure μ if for any $A \supseteq B \in \mathfrak{R}$ one has $\mu(B) \geq 0$; analogous with **negative**. Note that any measure space Ω may be partitioned into a disjoint union of a positive Ω^+ and a negative Ω^- part (**Hahn partition**). Any such partition allows the **Jordan decomposition** of the charge or measure into positive and negative parts (observe that the negative part is a *positive* measure) via $\forall A \in \mathfrak{R}$: $\mu^+(A) := \mu(A \cap \Omega^+)$ and $\mu^-(A) := -\mu(A \cap \Omega^-)$. Obviously this satisfies $\mu = \mu^+ - \mu^-$, and $|\mu| = \mu^+ + \mu^-$ is the **total variation** charge or measure. Then $|\mu|(\Omega)$ is called the total variation of μ. It is a norm on the spaces of charges and measures. In case the total variation is finite, the measure is called **bounded**. The space of bounded, signed charges is denoted by $\boldsymbol{ba}(\Omega, \mathfrak{R})$, and the space of bounded, signed measures by $\boldsymbol{ca}(\Omega, \mathfrak{A})$.

Both $\boldsymbol{ba}(\Omega, \mathfrak{R})$ and $\boldsymbol{ca}(\Omega, \mathfrak{A})$ are ordered vector space by saying that $\nu \leq \mu \Leftrightarrow \forall A \in \mathfrak{A} : \ \nu(A) \leq \mu(A)$. They are even Riesz spaces (see section 5). This means that the positive part μ^+, the negative part μ^-, and the absolute value $|\mu|$ may be defined, and they agree with what was just given in the classical fashion in the previous paragraph.

The **support** $\operatorname{supp} \mu$ of a measure μ on a measurable space $(\Omega, \mathfrak{A})$ is the smallest set $A \in \mathfrak{A}$ such that $\Omega \setminus A$ is a *null-set* for $|\mu|$. One also says that μ is **concentrated** on A. Two measures μ and ν are called **singular** or **orthogonal** to each other if their respective supports are disjoint, denoted by $\mu \perp \nu \Leftrightarrow \operatorname{supp} \mu \cap \operatorname{supp} \nu = \emptyset$. For any measure μ, in this sense $\mu^+ \perp \mu^-$.

For *integrable* $f \in L_0(\Omega)$, the *Lebesgue integral* is denoted by $\int_\Omega f \, \mathrm{d}\mu = \int_\Omega f(\omega) \, \mu(\mathrm{d}\omega)$. For a finite measure space ($\mu(\Omega) < \infty$), the space $L_0(\Omega)$ is **metrisable** in the following way. Define $d_\ell(f, g)(\omega)$ by either $|f(\omega) -$

$g(\omega)|/(1+|f(\omega)-g(\omega)|)$ or $\min\{|f(\omega)-g(\omega)|,1\}$, then a *translation invariant* metric or F-norm (see section 4.1) defining **convergence in measure** on $L_0(\Omega)$ is given by

$$d(f,g) := \int_\Omega d_\ell(f,g)(\omega)\,\mu(\mathrm{d}\omega).$$

Those functions where the integral of $|f|$ is finite are denoted by $L_1(\Omega,\mathfrak{A},\mu)$, or $L_1(\Omega,\mu)$ if the σ-algebra is clear from the context, or $L_1(\mu)$ if the basic space and σ-algebra are clear, or just $L_1(\Omega)$ if the measure is clear from the context. Such a function defines a new signed measure on Ω, denoted μf, via the assignment $\forall A \in \mathfrak{A} : \ \mu f(A) := \int_A f\,\mathrm{d}\mu = \int_A f(\omega)\,\mu(\mathrm{d}\omega)$. If $f \geq 0$ a.e., this is a *positive measure.*

A measure ν is called **absolutely continuous** w.r.t. the measure μ, denoted by $\nu \ll \mu$, iff $\forall A \in \mathfrak{A} : \ \mu(A) = 0 \Rightarrow \nu(A) = 0$. If we look at signed measures, these conditions have to hold for the *variation* measures $|\nu|$ and $|\mu|$. In that case $\exists g \in L_1(\Omega,\mu) : \ \nu(A) = \int_A g\,\mathrm{d}\mu$, and $g =: \mathrm{d}\nu/\mathrm{d}\mu$ is called the **Radon-Nikodým derivative**. Note that μf is absolutely continuous w.r.t. μ, and f is the Radon-Nikodým derivative. In case both $\nu \ll \mu$ and $\mu \ll \nu$, the measures are called **equivalent**, denoted by $\nu \simeq \mu$ (this is an equivalence relation). Remember the **Lebesgue decomposition**, in that for any two signed measures μ and ν there is a unique decomposition $\nu = \nu_c + \nu_s$, with $\nu_c \perp \nu_s$, where $\nu_c \ll \mu$ is the *absolutely continuous* part, and $\nu_s \perp \mu$ is the *singular* part.

3.4 Measures and Topology

Let $\mathfrak{A}$ be a σ-algebra on a topological space Ω with topology $\mathfrak{T}$, and μ a measure on $\mathfrak{A}$. Then the **outer measure** of μ for $A \in \mathfrak{A}$ is

$$\overline{\mu}(A) = \inf\{\mu(O) : \ A \subseteq O \in \mathfrak{T}\},$$

and the **inner measure** is

$$\underline{\mu}(A) = \sup\{\mu(K) : \ A \supseteq K \text{ compact}\}.$$

The measure is termed **outer regular** if $\forall A \in \mathfrak{A} : \ \mu(A) = \overline{\mu}(A)$, and **inner regular** or **tight** if $\mu(A) = \underline{\mu}(A)$; it is called **regular** if it is both inner and outer regular. The measure is **locally finite** if every point has a neighbourhood with finite measure.

In case Ω is a topological space with topology $\mathfrak{T}$, the σ-algebra generated by the open sets $\mathfrak{B} = \Sigma(\mathfrak{T})$ is called the **Borel algebra**. A measure μ on $\mathfrak{B}$ with $\mu(K) < \infty$ for any compact set $K \in \mathfrak{B}$ is called a **Borel measure**. The vector space of all signed, regular, and finite (total variation $|\mu|(\Omega) < \infty$) such Borel measures is denoted by $\mathscr{M}(\Omega) \subseteq \boldsymbol{ca}(\Omega, \mathfrak{B})$, with the total variation as norm this is a Banach space. The measurable functions w.r.t. $\mathfrak{B}$ are called Borel functions. For a Borel measure μ, the **topological support** $\operatorname{supp}_{\mathfrak{T}} \mu$ is the smallest closed set S such that $\mu(\Omega \setminus S) = 0$. If no confusion is possible with the support defined in the previous section 3.3, this is also just called the support.

The smallest σ-algebra such that all $\phi \in C_{00}(\Omega)$ are measurable is called the **Baire algebra** $\mathfrak{B}_0 = \Sigma(C_{00}(\Omega))$. Clearly $\mathfrak{B}_0 \subseteq \mathfrak{B}$. If Ω is locally compact, then $\mathfrak{B}_0$ is generated by the compact $\mathcal{G}_\delta$-sets. In this case, a measure on $\mathfrak{B}_0$ with finite values on compact sets (as before in the case of *Borel measures*) is called a **Baire measure**. The space of all such signed measures—they are automatically regular—is denoted by $\mathscr{M}_0(\Omega) \subseteq \boldsymbol{ca}(\Omega, \mathfrak{B}_0)$. Note that on a metric space $\mathfrak{B}_0 = \mathfrak{B}$, and hence Baire and Borel measures coincide.

The smallest σ-algebra such that all $\phi \in C_b(\Omega)$ (or $\phi \in C(\Omega)$) are measurable is called the **weak Baire algebra** $\mathfrak{B}_b = \Sigma(C_b(\Omega)) = \Sigma(C(\Omega))$. Clearly $\mathfrak{B}_0 \subseteq \mathfrak{B}_b \subseteq \mathfrak{B}$. A measure on $\mathfrak{B}_b$ with finite values on compact sets (as before in the case of *Borel measures*) is called a **weak Baire measure**. The space of all such signed, regular, and finite measures is denoted by $\mathscr{M}_b(\Omega) \subseteq \boldsymbol{ca}(\Omega, \mathfrak{B}_b)$.

A locally finite and inner regular *Borel measure* is called a **Radon measure**. It is a positive, continuous, linear functional on $C(\Omega)$ when Ω is locally compact. The space of all such signed measures is denoted by $\mathscr{M}_r(\Omega) \subseteq \boldsymbol{ca}(\Omega, \mathfrak{B})$. Note that on a complete, separable, metric space, every Borel measure is a Radon measure.

In case Ω is a topological group—see section 4.1, a measure which is invariant under the group action is called a **Hurwitz invariant integral** or **Haar measure**. It is essentially unique. **Lebesgue measure** on $\mathbb{R}$ is the complete *Borel measure* which satisfies $\mu([a, b]) = b - a$, and as it is translation invariant, it is also a *Haar measure*. Likewise, the corresponding *product measure* on $\mathbb{R}^n$ is also called *Lebesgue measure*.

The various spaces of measures are denoted by $\mathscr{M}(\Omega)$ with various qualifications (Baire measures, Borel measures, Radon measures, etc.). The natural duality pairing between a space of continuous functions and a space of measures is

$$f \in C(\Omega),\ \mu \in \mathscr{M}(\Omega) :\ \langle \mu, f \rangle := \int_\Omega f \,\mathrm{d}\mu = \int_\Omega f(\omega)\, \mu(\mathrm{d}\omega).$$

Another space of functions worth mentioning are the bounded measurable functions in general, not just for topological spaces, $\mathscr{B}_b(\Omega, \mathfrak{A}) = \mathscr{B}(\Omega) \cap$

$L_0(\Omega, \mathfrak{A})$. The charges $\boldsymbol{ba}(\Omega, \mathfrak{A})$ are in duality with this space of bounded measurable functions, the duality denoted in the same way.

We have the following representations for the dual space:

$$\begin{aligned}
\mathscr{B}_b(\Omega, \mathfrak{A})^* &\cong \boldsymbol{ba}(\Omega, \mathfrak{A}), \\
C_{00}(\Omega)^* &\cong \mathscr{M}_0(\Omega), \quad \Omega \text{ locally compact Hausdorff}, \\
C_0(\Omega)^* &\cong \mathscr{M}(\Omega), \quad \Omega \text{ locally compact Hausdorff}, \\
C_b(\Omega)^* &\cong \mathscr{M}_b(\Omega), \quad \Omega \text{ locally compact Hausdorff}, \\
C(\Omega)^* &\cong \mathscr{M}_r(\Omega), \quad \Omega \text{ locally compact Hausdorff}.
\end{aligned}$$

3.5 Probability Measures

In case $\mathbb{P} = \mu$ is a probability measure (i.e. $\mathbb{P}$ is positive and $\mathbb{P}(\Omega) = 1$), the measurable functions are called **random variables** (RVs), and the **expectation** of a RV $f \in L_1(\Omega, \mathfrak{A}, \mathbb{P})$ is

$$\bar{f} := \mathbb{E}(f) := \int_\Omega f \, \mathrm{d}\mathbb{P} = \int_\Omega f(\omega) \, \mathbb{P}(\mathrm{d}\omega).$$

Convergence in measure is in this case called **convergence in probability.** Associated to a random variable is its *law*, the measure $f_*\mathbb{P}$; the push-forward of the probability measure $\mathbb{P}$. The law is also a probability measure, but on the co-domain, or on $\operatorname{im} f$.

Assume Ω is a locally compact, σ-compact, and metrisable topological space with the σ-algebra of Borel sets $\mathfrak{B}$. The space of all signed bounded Borel measures $\mathscr{M}(\Omega, \mathfrak{B})$ is dual to the space $C_0(\Omega)$. It will be considered with the weak* topology $\sigma(\mathscr{M}(\Omega), C_0(\Omega))$. In case Ω is only locally compact, the duality is between $C_0(\Omega)$ and $\mathscr{M}_{reg}(\Omega, \mathfrak{B}_0)$, the subspace of signed, bounded, and regular Baire measures.

Whereas the **weak*** topology on $\mathscr{M}(\Omega, \mathfrak{B})$ is the weakest such that $f \mapsto \int_\Omega f \, \mathrm{d}\mu$ is continuous for each $\mu \in \mathscr{M}(\Omega, \mathfrak{B})$ and all $f \in C_0(\Omega)$), the **vague** topology is the weakest such that this is true for all $f \in C_{00}(\Omega)$), and the **weak** topology is the weakest such that this is true for all $f \in C_b(\Omega)$). There is possible confusion here with the functional analytic usage of the terms, as it only agrees for the weak* topology. The other topologies may be seen as induced by other duality pairings. Observe that the vague topology is the coarsest of these, and the weak* topology is coarser than the weak topology.

The cone of positive bounded Borel measures in $\mathscr{M}(\Omega, \mathfrak{B})$ has as a subset $\mathscr{P}(\Omega)$, the set of **probability measures** with

$$\forall \mathbb{P} \in \mathscr{P}(\Omega) : \ \mathbb{P}(\Omega) = \int_\Omega \chi_\Omega \, \mathrm{d}\mathbb{P} = \langle \mathbb{P}, \chi_\Omega \rangle = 1.$$

Convergence in the weak* topology is in this case also called convergence **in distribution**. The set Ω may be embedded into $\mathscr{P}(\Omega)$ via the evaluation duality, see section 1.6, which assigns to each $\omega \in \Omega$ the **point mass** or **δ-functional** δ_ω. For $A \in \mathfrak{B}$, one then has $\delta_\omega(A) = 1$ if $\omega \in A$, and $\delta_\omega(A) = 0$ otherwise.

4 Vector Spaces

This is an enlargement on recommendations by *Householder*. For an abstract vector space and its subsets, use capital letters in *calligraphic* font, e.g. $\mathcal{UVWXYZ}$. Denote simple elements of a vector space by lower case *Latin* letters, i.e. $x \in \mathcal{X}$. Denote scalars in this context preferably with lower case *Greek* letters, i.e. $\alpha \in \mathbb{K}$, and subsets by upper case *Greek* letters, i.e. $\Omega \subseteq \mathbb{K}$.

For a subset $\mathcal{A}$ of a vector space $\mathcal{X}$ over $\mathbb{K}$, the scalar multiples are defined for $\alpha \in \mathbb{K}$ by $\alpha\mathcal{A} := \{\alpha x \,|\, x \in \mathcal{A}\}$. Similarly, the sum of two subsets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{X}$ is $\mathcal{A} + \mathcal{B} = \{z = x + y \,|\, x \in \mathcal{A},\ y \in \mathcal{B}\}$.

4.1 Topological Vector Spaces

Start with a simpler situation for a moment: A group $\mathcal{G}$ with group operation $g \diamond h$, neutral element e, and inverse g^{-1}, which as a set also carries a topology $\mathfrak{T}_{\mathcal{G}}$ is a **topological group** iff the map $\mathcal{G} \times \mathcal{G} \ni (g, h) \mapsto g \diamond h^{-1} \in \mathcal{G}$ is continuous.

The neighbourhood filter $\mathfrak{U}(g)$ of any $g \in \mathcal{G}$ may be obtained from the neighbourhood filter $\mathfrak{U}(e)$ of e via $\mathfrak{U}(g) := g \diamond \mathfrak{U}(e) = \mathfrak{U}(e) \diamond g$. Hence the topology is in some sense invariant. For every $\mathcal{U} \in \mathfrak{U}(e)$ define the sets $\mathcal{L}_{\mathcal{U}} := \{(g, h) \in \mathcal{G} \times \mathcal{G} :\ g^{-1} \diamond h \in \mathcal{U}\}$ and $\mathcal{R}_{\mathcal{U}} := \{(g, h) \in \mathcal{G} \times \mathcal{G} :\ h \diamond g^{-1} \in \mathcal{U}\}$. Then the collections $\mathfrak{W}_L := \{\mathcal{L}_U :\ U \in \mathfrak{U}(e)\}$ and $\mathfrak{W}_R := \{\mathcal{R}_{\mathcal{U}} :\ \mathcal{U} \in \mathfrak{U}(e)\}$ are bases for two *uniformities*; obviously they coincide in case $\mathcal{G}$ is an Abelian group. In any case, every topological group has a uniform topology. For any $\mathcal{U} \in \mathfrak{U}(e)$ the mapping $g \mapsto \mathcal{U}(g) := g \diamond \mathcal{U}$ defines a *covering system*, and all such covering systems are the *neighbourhood* system of the uniformity.

A vector space $\mathcal{X}$ is a **topological vector space** (TVS) if the (Abelian) additive group $\mathcal{X}$ is a topological group, and if additionally scalar multiplication $\mathbb{K} \times \mathcal{X} \ni (\alpha, x) \mapsto \alpha x \in \mathcal{X}$ is continuous. Hence any topological vector space has a uniform topology. As the topology is invariant under the group operation, it is more specifically *translation invariant.* Therefore the pseudo-metrics in the gauge system generating the uniformity are also translation invariant. Likewise, if the topology is metrisable, the metric is translation invariant.

Hence these pseudo-metrics or metrics $d(x, y)$ can be investigated at $y = 0 \in \mathcal{X}$, therefore consider $q(x) := d(x, 0)$. The functions q satisfy $\forall \alpha \in \mathbb{K}, |\alpha| \leq 1 : \; q(\alpha x) \leq q(x)$, and $\lim_{n\to\infty} q(x/n) = 0$, as well as the triangle inequality $q(x + y) \leq q(x) + q(y)$. Such functions are called **F-seminorms**. They define a sub-basis for the zero-neighbourhood-filter—open 'balls'—via

$$B_q(0, \varepsilon) := \mathcal{U}_{q,\varepsilon}(0) := \{x \in \mathcal{X} \,|\, q(x) < \varepsilon\} \in \mathfrak{U}(0).$$

In case the F-seminorm is definite $((q(x) = 0) \Rightarrow (x = 0))$ it is called a **F-norm**. Hence the topology and uniformity of any TVS may be generated by a collection (gauge system) of F-seminorms, namely exactly the *continuous* F-seminorms. In case the topology is metrisable (generated by a countable number of F-seminorms), the topology and uniformity may actually be generated by a F-norm. In any case, if the underlying metrisable uniform space is complete, it is called an **F-space**.

A subset $\mathcal{A} \subseteq \mathcal{X}$ of a TVS is called **bounded**, if for any zero-neighbourhood $\mathcal{U} \in \mathfrak{U}(0)$ there is a $\lambda > 0$ such that $\lambda\,\mathcal{A} \subseteq \mathcal{U}$. This means equivalently that for any continuous F-seminorm q one has $\sup\{q(x) \,|\, x \in \mathcal{A}\} < \infty$.

If a F-seminorm q is positively homogeneous and symmetric—or circular in the complex case, i.e. it satisfies $\forall \alpha \in \mathbb{K} : \; q(\alpha x) = |\alpha| q(x)$, it is called a **seminorm**, and in the definite case it is called a **norm** and denoted by $\|\cdot\|_{\mathcal{X}}$ or just $\|\cdot\|$ if the norm is clear from the context. In both cases it is a convex function(al), see section 12. Sometimes norms may also be denoted by $|\!|\!|\cdot|\!|\!|$, e.g. for norms of linear maps, or by $|\cdot|$, where the latter notation is sometimes also used to denote a seminorm, and also the absolute value in ordered vector spaces—see section 5; it should be clear from the context what is meant. Note that for a seminorm the set $\{x : \; q(x) < \varepsilon\}$ is convex. This means that the open balls $B_{\|\cdot\|}(0, \varepsilon)$ are absolutely convex sets.

If the topology and uniformity is generated by a collection of seminorms, the TVS is a **locally convex space** (LCS), i.e. every point has a neighbourhood basis of convex sets. In case the space is also metrisable and complete, it is called a **Fréchet space**. If the topology and uniformity is generated by a norm, $\mathcal{X}$ is a **normed** space, and if complete is called a **Banach space**.

In case the norm comes from an **inner product** $\langle\cdot|\cdot\rangle$ with $\|x\| = \sqrt{\langle x|x\rangle}$, the space is called a inner product space or pre-Hilbert space. If the space is complete, it is called a **Hilbert space**. Note that the norm derived from an inner product satisfies the **parallelogram identity**—this is a defining property of such a norm: $\|x + y\|^2 + \|x - y\|^2 = 2(\|x\|^2 + \|y\|^2)$.

Apart from the F-norms which are sub-additive but not necessarily positively homogeneous, one may use functionals $q : \mathcal{X} \to \mathbb{R}_+$ which satisfy all the conditions for a seminorm, i.e. are positively homogeneous, but not necessarily sub-additive. If instead they satisfy for some $K > 0$

$$q(x+y) \leq K(q(x)+q(y)),$$

they are called **quasi-seminorms**, and in case they are definite **quasi-norms**. As before one defines the open 'balls' $B_q(0,\varepsilon)$ as a sub-basis for the zero-neighbourhood-filter. This may be used to define the uniform structure on $\mathcal{X}$. If the resulting space is complete, it is called a **quasi-Banach space**.

4.2 Constructions with Vector Spaces

If $\{\mathcal{X}_j\}_{j\in J}$ is a family of vector spaces, then the **product** $\mathcal{X} = \prod_{j\in J}\mathcal{X}_j$ is naturally a vector space under componentwise operations, for finitely many spaces this is also denoted as $\mathcal{X} = \mathcal{X}_1 \times \cdots \times \mathcal{X}_n$. Another construction is the **external direct sum** $\mathcal{Y} = \bigoplus_{j\in J}\mathcal{X}_j$, also denoted as $\sum_{j\in J}\mathcal{X}_j$, where for each $\mathcal{Y} \ni y = \sum_{j\in J} x_j$ with $x_j \in \mathcal{X}_j$ only finitely many terms are non-zero. In case all $\mathcal{X}_j$ are subspaces of some space $\mathcal{X}$, one may form the **internal direct sum** $\mathcal{Z} = \sum_{j\in J}\mathcal{X}_j$. In case there is one unique way of writing each $\mathcal{Z} \ni z = \sum_{j\in J} x_j$ —this is the case if $\forall j,k \in J : j \neq k \Rightarrow \mathcal{X}_j \cap \mathcal{X}_k = \{0\}$ —then this is denoted as $\mathcal{Z} = \bigoplus_{j\in J}\mathcal{X}_j$. In this case the somewhat opposite operation is $\mathcal{Z} \ominus \mathcal{X}_k = \bigoplus_{j\neq k}\mathcal{X}_j$. In section 7 these constructions are considered together with possible topologies on the spaces $\mathcal{X}_j$, and on the products and sums.

If $\mathcal{X}$ is a vector space and J any set, then $\mathcal{X}^J = \prod_{j\in J}\mathcal{X}$ is naturally a vector space—the set of all maps $\mathscr{F}(J,\mathcal{X})$ —under pointwise definitions of the algebraic operations, as is the subspace $\mathcal{X}^{(J)} = \bigoplus_{j\in J}\mathcal{X}$, where $x_j = 0$ except for finitely many j. Remember again that $(\mathcal{X}^J)^* \cong (\mathcal{X}^*)^{(J)}$ with the natural duality pairing

$$\forall x = (\xi_j) \in \mathcal{X}^J, y = (\eta_j) \in (\mathcal{X}^*)^{(J)} :\ \langle x, y\rangle := \sum_{j\in J}\langle \xi_j, \eta_j\rangle_{\mathcal{X}\times\mathcal{X}^*},$$

and especially that $\boldsymbol{\omega}^* = (\mathbb{R}^{\mathbb{N}})^* \cong \mathbb{R}^{(\mathbb{N})} = \boldsymbol{c}_{00}$.

If $\mathcal{X}$ is a LCS and $\mathcal{Y}$ a closed subspace, then the quotient space is denoted by $\mathcal{X}/\mathcal{Y}$. The dual of this space is isomorphic to the "orthogonal complement"

$$(\mathcal{X}/\mathcal{Y})^* \cong \mathcal{Y}^\perp := \{y^* \in \mathcal{X}^* \,|\, \forall y \in \mathcal{Y} :\ \langle y, y^*\rangle = 0\} \subseteq \mathcal{X}^*.$$

Note that for any subspace $\mathcal{Y} \subseteq \mathcal{X}$ one has $\mathcal{Y}^o = \mathcal{Y}^\perp$, and ${}^\perp(\mathcal{Y}^\perp) = \operatorname{cl}\mathcal{Y}$.

4.3 Products, Sums, and Limits

Let $\{\mathcal{X}_j\}_{j\in J}$ be a family of TVS and $\mathcal{X}$ a vector space. Assume that for each $j \in J$ a linear map $T_j : \mathcal{X} \to \mathcal{X}_j$ is given. The *coarsest* topology on $\mathcal{X}$ which makes all T_j continuous makes $\mathcal{X}$ into a TVS, and is called the **projective** or **initial** topology. The product $\prod_{j\in J} \mathcal{X}_j$ together with the natural projections carries such a topology, here also called the **product topology**. For a **projective system** we require additionally that J is a directed set, and that $\forall (j,k) \in J \times J$ such that $j \preceq k$, there is a continuous linear map $T_{jk} : \mathcal{X}_j \to \mathcal{X}_k$ with dense image $\mathrm{cl}(\mathrm{im}\, T_{jk}) = \mathcal{X}_k$ such that $T_{jj} = I_{\mathcal{X}_j}$, and $T_{ik} = T_{ij} \circ T_{jk}$ whenever $i \preceq j \preceq k$. Then the subspace of $\prod_{j\in J} \mathcal{X}_j$ where $T_{jk} x_k = x_j$ holds whenever $j \preceq k$ is called the **projective limit** $\mathrm{proj}\lim_{j\in J} \mathcal{X}_j$ of the family.

Conversely assume that $\{\mathcal{X}_j\}_{j\in J}$ is a family of TVS and $\mathcal{X}$ a vector space, with injective linear maps $S_j : \mathcal{X}_j \to \mathcal{X}$. The **finest** topology on $\mathcal{X}$ which makes all S_j continuous makes $\mathcal{X}$ into a TVS, and is called the **injective** or **final** topology. The sum $\bigoplus_{j\in J} \mathcal{X}_j$ together with the natural injections carries such a topology. For an **injective system** we require additionally that J is a directed set, and that $\forall (j,k) \in J \times J$ such that $j \preceq k$, there is a continuous injective linear map $S_{kj} : \mathcal{X}_j \to \mathcal{X}_k$ such that $S_{jj} = I_{\mathcal{X}_j}$, and $S_{ki} = S_{kj} \circ S_{ji}$ whenever $i \preceq j \preceq k$. Let $\iota_j : \mathcal{X}_j \to \bigoplus_{j\in J} \mathcal{X}_j$ be the natural injection, and $\mathcal{L} = \mathrm{span}\left(\bigcup_{i\preceq j} \mathrm{im}(\iota_i - \iota_j \circ S_{ji})\right)$. Then $\bigoplus_{j\in J} \mathcal{X}_j / \mathcal{L}$ is called the **injective limit** $\mathrm{inj}\lim_{j\in J} \mathcal{X}_j$ of the family.

Assume now a space $\mathcal{H}$ equipped with a countable family of **compatible** inner products $\{\langle x, y\rangle_j\}_{j\in\mathbb{N}}$. Compatible means that for a sequence $\{x_n\}_{n\in\mathbb{N}}$ which converges in the norm of one inner product ($\|x_n - x\|_j \to 0$) and is *Cauchy* in another ($\forall \varepsilon > 0\ \exists n_0 \in \mathbb{N},\ n, m \geq n_0 :\ \|x_n - x_m\|_k \leq \varepsilon$), it holds that also $\|x_n - x\|_k \to 0$. Such a space is called **countably Hilbert**. Suppose further that $\forall x \in \mathcal{H} :\ \|x\|_j \leq \|x\|_k$ for $j \leq k$. Then we have $\mathcal{H}_k \hookrightarrow \mathcal{H}_j$, where $\mathcal{H}_j$ is the completion of $\mathcal{H}$ in the norm $\|\cdot\|_j$. Let $\mathcal{K} = \bigcap_{j\in\mathbb{N}} \mathcal{H}_j$ with the *projective limit topology*, a *Fréchet space*. If for every $j \in \mathbb{N}$ there is a $k \in \mathbb{N}$ such that the inclusion $\iota_{kj} : \mathcal{H}_k \hookrightarrow \mathcal{H}_j$ is *nuclear* (or, essentially equivalent, *Hilbert-Schmidt*)—see setion 4.6.1—then $\mathcal{K}$ is called a **nuclear** space. Let $\mathcal{H}_{-j} = \mathcal{H}_j^*$, then $\mathcal{K}^* \cong \bigcup_{j\in\mathbb{N}} \mathcal{H}_{-j}$ with the *injective limit topology*.

4.4 Linear and Multi-Linear Maps

Linear mappings (operators) on a vector space $\mathcal{X}$ can be denoted by upper case *Latin* letters, i.e. $\forall x \in \mathcal{X},\ y = Ax \in \mathcal{Y}$. For locally convex spaces (LCS) $\mathcal{X}$ and $\mathcal{Y}$, the space of all continuous linear maps from $\mathcal{X}$ to $\mathcal{Y}$ is denoted by $\mathscr{L}(\mathcal{X}, \mathcal{Y}) \subseteq C_u(\mathcal{X}, \mathcal{Y}) \subseteq C(\mathcal{X}, \mathcal{Y}) \subseteq \mathscr{F}(\mathcal{X}, \mathcal{Y})$. For $\mathscr{L}(\mathcal{X}, \mathcal{X})$ write $\mathscr{L}(\mathcal{X})$.

The identity is denoted by $I \in \mathscr{L}(\mathcal{X})$. In case they have to be considered, the space of *all* linear maps—continuous or not—may be denoted by $\mathrm{L}(\mathcal{X}, \mathcal{Y})$. Clearly $\mathscr{L}(\mathcal{X}, \mathcal{Y}) \subseteq \mathrm{L}(\mathcal{X}, \mathcal{Y})$, with equality in case $\mathcal{X}$ is finite dimensional. Remember that a linear map A on a LCS is continuous if it is continuous at zero, equivalently if it is bounded (maps bounded sets into bounded sets), or in case the spaces $\mathcal{X}, \mathcal{Y}$ are *Fréchet spaces* (see section 4.1), according to the *closed graph theorem*, if the graph of A (see section 1.6),

$$\operatorname{gra} A := \{(x, y) : \ x \in \mathcal{X},\, y = Ax\}$$

is closed as a subspace of $\mathcal{X} \times \mathcal{Y} \cong \mathcal{X} \oplus \mathcal{Y}$.

In case the spaces $\mathcal{X}, \mathcal{Y}$ are *normed spaces*, so is $\mathscr{L}(\mathcal{X}, \mathcal{Y})$, by defining for $A \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ the **operator-norm**

$$\|A\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} := \sup_{x \neq 0} \frac{\|Ax\|_{\mathcal{Y}}}{\|x\|_{\mathcal{X}}}.$$

Observe that for $A \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ and $B \in \mathscr{L}(\mathcal{Y}, \mathcal{Z})$ it holds that

$$\|B \circ A\|_{\mathscr{L}(\mathcal{X},\mathcal{Z})} \leq \|A\|_{\mathscr{L}(\mathcal{X},\mathcal{Y})} \|B\|_{\mathscr{L}(\mathcal{Y},\mathcal{Z})}.$$

The space $\mathscr{L}(\mathcal{X}, \mathcal{X}) = \mathscr{L}(\mathcal{X})$ with this norm becomes a **normed algebra**. In case $\mathcal{Y}$ is a *Banach space*, so is $\mathscr{L}(\mathcal{X}, \mathcal{Y})$, and hence $\mathscr{L}(\mathcal{X})$ is a **Banach algebra**.

Composition of maps $A, B \in \mathscr{L}(\mathcal{X})$, i.e. $AB := A \circ B$, makes $\mathscr{L}(\mathcal{X})$ into an **associative algebra** with unit. If for some $n \in \mathbb{N}$ one has that $A^n = 0$, the map $A \in \mathscr{L}(\mathcal{X})$ is called **nilpotent.** The commutator of two maps is defined like for any associative algebra by $[A, B] := A \circ B - B \circ A = AB - BA$, and this as usual makes $\mathscr{L}(\mathcal{X})$ then into a *Lie* algebra.

For a linear operator $A \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$, the **kernel** or **null space** is

$$\ker A = \{x \in \mathcal{X} | Ax = 0 \in \mathcal{Y}\} = A^{-1}(\{0\}) \subseteq \mathcal{X},$$

and the **co-kernel** is $\operatorname{coker} A := \mathcal{Y} / \operatorname{im} A$.

In a LCS, particular subspaces of $\mathscr{L}(\mathcal{X}, \mathcal{Y})$ are the **compact** operators $\mathscr{L}_0(\mathcal{X}, \mathcal{Y}) \subseteq \mathscr{L}(\mathcal{X}, \mathcal{Y})$ (mapping bounded sets in $\mathcal{X}$ into relatively compact sets in $\mathcal{Y}$), and the operators of **finite rank** $\mathscr{L}_{00}(\mathcal{X}, \mathcal{Y}) \subseteq \mathscr{L}_0(\mathcal{X}, \mathcal{Y})$, i.e. for $A \in \mathscr{L}_{00}(\mathcal{X}, \mathcal{Y})$ it holds that $\operatorname{rank}(A) := \dim(\operatorname{im} A) < \infty$. If both $\dim(\ker A) < \infty$ and $\dim(\operatorname{coker} A) < \infty$, then $\operatorname{ind}(A) := \dim(\ker A) - \dim(\operatorname{coker} A)$ is the index of $A \in \mathscr{L}_\Phi(\mathcal{X}, \mathcal{Y})$, the space of **Fredholm** operators.

Multi-linear maps are denoted by

$$\mathscr{L}^k(\mathcal{X}_1, \cdots, \mathcal{X}_k; \mathcal{Y}) := \mathscr{L}(\mathcal{X}_1 \times \cdots \times \mathcal{X}_k, \mathcal{Y})$$

for a k-linear map, and especially if the $\mathcal{X}_\jmath$ are all equal just with $\mathscr{L}^k(\mathcal{X};\mathcal{Y}) := \mathscr{L}(\mathcal{X}^k,\mathcal{Y})$, and if $\mathcal{Y} = \mathcal{X}$ just with $\mathscr{L}^k(\mathcal{X})$. Continuous multi-linear maps again map bounded subsets of $\mathcal{X}^k$ into bounded subsets of $\mathcal{Y}$. In case the spaces $\mathcal{X},\mathcal{Y}$ are *normed spaces*, so is $\mathscr{L}^k(\mathcal{X};\mathcal{Y})$, but the norm is not canonically given; one example for $B \in \mathscr{L}^k(\mathcal{X};\mathcal{Y})$ is

$$\|B\|_{\mathscr{L}^k(\mathcal{X};\mathcal{Y})} := \sup\{\|Bx\|_\mathcal{Y} \,:\, x = (x_1,\dots,x_k) \in \mathcal{X}^k,\ \sum_{m=1}^{k} \|x_m\|_\mathcal{X} \le 1\}.$$

Let $\mathfrak{S}_k$ be the symmetric group of order k (group of all permutations ϖ of k elements). Each $\varpi \in \mathfrak{S}_k$ generates a map in $\mathscr{L}(\mathscr{L}^k(\mathcal{X};\mathcal{Y}))$. For $B \in \mathscr{L}^k(\mathcal{X};\mathcal{Y})$ and $\varpi \in \mathfrak{S}_k$, define $\varpi B \in \mathscr{L}^k(\mathcal{X};\mathcal{Y})$ by

$$\varpi B(x_1,\dots,x_k) := B(x_{\varpi(1)},\dots,x_{\varpi(k)}).$$

The **symmetriser** $S_k \in \mathscr{L}(\mathscr{L}^k(\mathcal{X};\mathcal{Y}))$ is for a $B \in \mathscr{L}^k(\mathcal{X};\mathcal{Y})$ defined by

$$S_k(B) := \frac{1}{k!} \sum_{\varpi\in\mathfrak{S}_k} \varpi B,$$

The resulting expression is symmetric in the k arguments. Similarly, define the **anti-symmetriser** or **alternator** $A_k \in \mathscr{L}(\mathscr{L}^k(\mathcal{X};\mathcal{Y}))$ for a $B \in \mathscr{L}^k(\mathcal{X};\mathcal{Y})$ by

$$A_k(B) := \frac{1}{k!} \sum_{\varpi\in\mathfrak{S}_k} \operatorname{sgn}(\varpi)\, \varpi B,$$

where $\operatorname{sgn}(\varpi)$ is the sign of the permutation ϖ. The resulting expression is anti-symmetric resp. alternating in the k arguments.

Both S_k and A_k are projectors, on the subspace of **symmetric** multi-linear maps in the first case, and on the subspace of **alternating** or **anti-symmetric** multi-linear maps in the latter.

4.5 Duality and Orthogonality

Special notation for the **dual** of a TVS $\mathcal{X}^* := \mathscr{L}(\mathcal{X},\mathbb{K})$. This is the **topological dual**, whereas the **algebraic dual** of all linear functionals—continuous or not—is denoted by $\mathcal{X}' := \mathrm{L}(\mathcal{X},\mathbb{K})$. Unfortunately, sometimes these notations are just the reverse.

If $x^* \in \mathcal{X}^*$, then $\forall x \in \mathcal{X} :\ \langle x^*, x\rangle := x^*(x) \in \mathbb{K}$ is the duality pairing. A subset $\mathcal{G} \subseteq \mathcal{X}^*$ is said to **separate the points**, if for each $x \in \mathcal{X}$ there is a $g \in \mathcal{G}$ such that $\langle g, x\rangle \neq 0$. Note that for a LCS (not necessarily for every TVS), $\mathcal{X}^* \neq \emptyset$ and separates the points.

For a sequence (or net) converging in the *(strong) topology* of a LCS $\mathcal{X}$, write $x_j \to x = \lim x_j$. In a normed space—the norm of an element $x \in \mathcal{X}$ denoted by $\|x\|$ —this means $\|x_j - x\| \to 0$. This is called **strong convergence**. For convergence in the *weak topology*, write $x_j \rightharpoonup x = \text{w-lim } x_j$, i.e. $\forall x^* \in \mathcal{X}^* : \langle x^*, x_j \rangle \to \langle x^*, x \rangle$, the topology is generated by the seminorms $|\langle x^*, \cdot \rangle|$. This is called **weak convergence**. For convergence in the *weak* topology* on $\mathcal{X}^*$, write $x_j^* \overset{*}{\rightharpoonup} x^* = \text{w*-lim } x_j^*$, i.e. $\forall x \in \mathcal{X} : \langle x_j^*, x \rangle \to \langle x^*, x \rangle$, the topology is generated by the seminorms $|\langle \cdot, x \rangle|$. This is called **weak* convergence**.

Various toplogies are defined by seminorms via certain subsets $\mathcal{A} \in \mathcal{X}^*$ as $q_{\mathcal{A}}(x) := sup_{x^* \in \mathcal{A}} |\langle x^*, x \rangle|$. Convergence in such a seminorm is equivalent to *uniform* convergence on the subset $\mathcal{A}$. The **strong topology** on a LCS $\mathcal{X}$ is denoted by $\beta(\mathcal{X}, \mathcal{X}^*)$, and is given by uniform convergence on all *bounded* subsets of $\mathcal{X}^*$. Likewise the *weak* topology is given by uniform convergence on all *finite* subsets of $\mathcal{X}^*$, and is denoted by $\sigma(\mathcal{X}, \mathcal{X}^*)$. The **Mackey topology** is denoted by $\tau(\mathcal{X}, \mathcal{X}^*)$, and is given by uniform convergence on all compact and absolutely convex subsets of $\mathcal{X}^*$. The **weak* topology** on the dual is denoted by $\sigma(\mathcal{X}^*, \mathcal{X})$ and is given by uniform convergence on all *finite* subsets of $\mathcal{X}$. Similarly one defines the *strong* topolgy on the dual $\beta(\mathcal{X}^*, \mathcal{X})$, and likewise the Mackey toplogy on the dual $\tau(\mathcal{X}^*, \mathcal{X})$. In case $\mathcal{X}$ is a normed space, $\mathcal{X}^*$ is a Banach space (because $\mathbb{K}$ is), with the strong topology given by the norm $\|x^*\|_{\mathcal{X}^*} := \sup_{\|x\|_{\mathcal{X}}=1} |\langle x^*, x \rangle|$.

Convergence of a sequence $A_j \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ to A in the *operator norm* is denoted by $A_j \Rightarrow A$, i.e. if $\|A - A_j\|_{\mathscr{L}} \to 0$, and is called **operator norm convergence**. The **strong operator convergence** $A_j \to A$ is defined as $\forall x \in \mathcal{X} : A_j x \to Ax$, the topology is generated by the seminorms $\| \cdot x\|$. Similarly **weak operator convergence** $A_j \rightharpoonup A$ is defined as $\forall x \in \mathcal{X} : A_j x \rightharpoonup Ax$, the topology is generated by the seminorms $|\langle x^*, \cdot x \rangle|$. A map $A \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ with $\|Ax\|_{\mathcal{Y}} = \|x\|_{\mathcal{X}}$ is called an **isometry**. Note that isometries are injective, but not necessarily surjective.

For $x, y \in \mathcal{X}$, the **line segment** joining them is $[x, y] := \{z = (1 - \vartheta)x + \vartheta y |\ \vartheta \in [0, 1]\}$. A subset $\mathcal{C} \subseteq \mathcal{X}$ is convex, if for any $x, y \in \mathcal{C}$ the *line segment* joining them is also in $\mathcal{C}$ ($[x, y] \subseteq \mathcal{C}$). For a subset $\mathcal{A} \subseteq \mathcal{X}$ of a LCS, the **convex hull** is denoted by $\operatorname{co} \mathcal{A}$—the smallest convex set containing $\mathcal{A}$, or equivalently the intersection of all such sets—and the closure of this set by $\overline{\operatorname{co}}\, \mathcal{A}$. The **absolutely convex hull** is denoted by $\operatorname{acx} \mathcal{A}$, and the closure of this by $\overline{\operatorname{acx}}\, \mathcal{A}$. The set of **extreme points** of a convex set $\mathcal{A}$ is denoted by $\operatorname{ext} \mathcal{A}$.

The linear hull or **span** of a set $\mathcal{A} \subseteq \mathcal{X}$ is denoted by $\operatorname{span} \mathcal{A}$, and the closure of the linear hull by $\overline{\operatorname{span}}\, \mathcal{A}$. Note that $\operatorname{span} \emptyset = \{0\}$, and that finite dimensional subspaces are always closed.

The **orthogonal complement** or **annihilator** of a set $\mathcal{A} \subseteq \mathcal{X}$ is the subspace

$$\mathcal{A}^{\perp} := \{x^* \in \mathcal{X}^* | \, \forall x \in \mathcal{A} : \ \langle x^*, x \rangle = 0\} \subseteq \mathcal{X}^*,$$

and for $\mathcal{B} \subseteq \mathcal{X}^*$ the subspace

$$\mathcal{B}^{\perp} := \{x \in \mathcal{X} | \, \forall x^* \in \mathcal{B} : \ \langle x^*, x \rangle = 0\} \subseteq \mathcal{X}.$$

Note that $\overline{\operatorname{span}}\, \mathcal{A} = \mathcal{A}^{\perp\perp}$. If $\langle x^*, x \rangle = 0$, one also writes $x^* \perp x$, and similarly for subsets $\mathcal{B} \perp \mathcal{A}$, if $\forall x^* \in \mathcal{B}, x \in \mathcal{A} : \ \langle x^*, x \rangle = 0$.

For $A \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ the **transpose** or **dual operator** $A^* \in \mathscr{L}(\mathcal{Y}^*, \mathcal{X}^*)$ is defined by

$$\forall x \in \mathcal{X} \ y^* \in \mathcal{Y}^* : \quad \langle A^* y^*, x \rangle = \langle y^*, Ax \rangle.$$

The transpose has the following properties:

$$\begin{aligned} A, B \in \mathscr{L}(\mathcal{X}, \mathcal{Y}) : \quad (A + B)^* &= \alpha A^* + B^*, \\ \alpha \in \mathbb{R} : \quad (\alpha A)^* &= \alpha A^*, \\ A \in \mathscr{L}(\mathcal{X}, \mathcal{Y}), B \in \mathscr{L}(\mathcal{Y}, \mathcal{Z}) : \quad (AB)^* &= B^* A^*, \\ A \in GL(\mathcal{X}, \mathcal{Y}) : \quad A^{-*} := (A^{-1})^* &= (A^*)^{-1}. \end{aligned}$$

If $\mathcal{X}$ and $\mathcal{Y}$ are just LCS, we have

$$(\operatorname{im} A)^{\perp} = \ker A^* \ \text{and} \ \operatorname{cl} \operatorname{im} A^* \subseteq \ker A^{\perp}.$$

For Banach spaces $\mathcal{X}$ and $\mathcal{Y}$ one has $\|A\| = \|A^*\|$. If in addition $\operatorname{im} A$ is closed, then also $\operatorname{im} A^*$ is closed and

$$\begin{aligned} \exists C > 0 \ \forall y \in \operatorname{im} A \ \exists x \in \mathcal{X} : \quad Ax &= y \ \text{and} \ \|x\| \le C \|y\| \\ \operatorname{im} A^* &= (\ker A)^{\perp}. \end{aligned}$$

If additionally the spaces are reflexive, then $(\mathcal{X}^*)^* \cong \mathcal{X}$, and one has

$$\begin{aligned} (A^*)^* &= A, & \operatorname{cl} \operatorname{im} A &= (\ker A^*)^{\perp}, \\ (\operatorname{im} A^*)^{\perp} &= \ker A, \quad \text{and} & \operatorname{cl} \operatorname{im} A^* &= (\ker A)^{\perp}. \end{aligned}$$

A **cone** $\mathcal{C}$ is a subset of a vector space $\mathcal{X}$, such that $\forall \lambda > 0 : \ \lambda \mathcal{C} \subseteq \mathcal{C}$. If $0 \in \mathcal{C}$, it is a **pointed cone**, otherwise a **blunt cone**. It is a **salient cone** if $\mathcal{C} \cap (-\mathcal{C}) \subseteq \{0\}$.

For any $\mathcal{A} \subseteq \mathcal{X}$, the **spanned cone** or **generated cone** is

$$\mathcal{A}^{\vee} := \bigcup_{\lambda \ge 0} \lambda \mathcal{A} \subseteq \mathcal{X}.$$

Note that this cone is convex iff $\mathcal{A}$ is so. The **recession cone** is

$$\mathcal{A}^\wedge := \bigcap_{\lambda \geq 0} \lambda \mathcal{A} \subseteq \mathcal{A}^\vee,$$

and the **barrier cone** of a set $\mathcal{A} \subseteq \mathcal{X}$ is

$$\mathcal{A}^\infty := \{x \in \mathcal{X} | \forall x^* \in \mathcal{B} : \ |\langle x^*, x \rangle| \leq \infty\} \subseteq \mathcal{X}.$$

The **absolute polar** (or just polar) of a set $\mathcal{A} \subseteq \mathcal{X}$ is the closed convex set

$$\mathcal{A}^o := \{x^* \in \mathcal{X}^* | \forall x \in \mathcal{A} : \ |\langle x^*, x \rangle| \leq 1\} \subseteq \mathcal{X}^*,$$

the **lower polar** is the closed convex set

$$\mathcal{A}^\flat := \{x^* \in \mathcal{X}^* | \forall x \in \mathcal{A} : \ \langle x^*, x \rangle \geq -1\} \subseteq \mathcal{X}^*,$$

and the **upper polar** is the closed convex set

$$\mathcal{A}^\sharp := \{x^* \in \mathcal{X}^* | \forall x \in \mathcal{A} : \ \langle x^*, x \rangle \leq 1\} \subseteq \mathcal{X}^*.$$

The **positive polar cone** of a set $\mathcal{A}$ is the closed convex set

$$\mathcal{A}^\oplus := \{x^* \in \mathcal{X}^* | \forall x \in \mathcal{A} : \ \langle x^*, x \rangle \geq 0\} \subseteq \mathcal{X}^*,$$

and the **negative polar cone** of a set $\mathcal{A}$ is the closed convex set

$$\mathcal{A}^\ominus := \{x^* \in \mathcal{X}^* | \forall x \in \mathcal{A} : \ \langle x^*, x \rangle \leq 0\} \subseteq \mathcal{X}^*.$$

The **bipolar** is $\mathcal{A}^{oo} := (\mathcal{A}^o)^o$. Note that $\mathcal{A}^{oo} = \overline{\text{acx}}\, \mathcal{A}$ and $\mathcal{A}^{\sharp\sharp} = \overline{\text{co}}\, (\mathcal{A} \cup \{0\})$, and that $\mathcal{A}^{\ominus\ominus} = \overline{\text{co}}\, \mathcal{A}$.

4.6 Reflexive Spaces

If $\mathcal{X} \cong (\mathcal{X}^*)^*$, where $\mathcal{X}^*$ carries the strong $\beta(\mathcal{X}^*, \mathcal{X})$ toplogy, the space $\mathcal{X}$ is called **reflexive**. On a reflexive space $\mathcal{X}$, for a linear operator $A \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ one has $A^* \in \mathscr{L}((\mathcal{X}^*)^*, \mathcal{X}^*)$, which then is viewed as $A^* \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$. If for $A \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ it holds that $A = A^*$, then A is called **symmetric** or **self-adjoint**. The symmetric operators are a subspace $\text{sym}(\mathcal{X}) \subset \mathscr{L}(\mathcal{X})$. If it holds that $A = -A^*$, then A is called **skew**, **skew-symmetric**, or **skew-adjoint**. This subspace is denoted by $\text{so}(\mathcal{X}) \subset \mathscr{L}(\mathcal{X})$.

A self-adjoint operator $A \in \mathscr{L}(\mathcal{X}, \mathcal{X}^*)$ is **positive (semi-definite)** if $\forall x \in \mathcal{X} : \langle Ax, x \rangle \geq 0$, and **positive definite** in case $\forall x \neq 0 : \ \langle Ax, x \rangle > 0$. These operators are a **salient, closed, convex cone** in $\mathscr{L}(\mathcal{X}, \mathcal{X}^*)$. Hence this defines a **partial order** on the subspace of self-adjoint operators, such that $M \leq A \Leftrightarrow A - M$ is positive, see section 1.5.

4.6.1 Inner Products

If $\mathcal{H}$ is a **Hilbert** space, the **inner product** of $x, y \in \mathcal{H}$ may be denoted by $\langle x|y\rangle$ or $(x|y)$. If $\mathcal{H}$ is identified with its dual, it may also be denoted as a duality pairing $\langle x, y\rangle$. In analogy to section 4.5 we write $x \perp y$ in case $\langle x|y\rangle = 0$. Similarly for subsets $\mathcal{A}, \mathcal{B} \subseteq \mathcal{H}$ one writes $\mathcal{A}^\perp$ and $\mathcal{A} \perp \mathcal{B}$, as well as $\mathcal{A}^o$, where the inner product as taken instead of the duality pairing. If in the Hilbert space context one has $\mathcal{X} \oplus \mathcal{Y} = \mathcal{Z}$ for some subspaces $\mathcal{X}, \mathcal{Y}, \mathcal{Z} \subseteq \mathcal{H}$, it is usually implied that $\mathcal{X} \perp \mathcal{Y}$. In case $\mathcal{X} \subset \mathcal{Y}$, the orthogonal complement of $\mathcal{X}$ in $\mathcal{Y}$ is $\mathcal{Y} \ominus \mathcal{X} := \mathcal{X}^\perp \cap \mathcal{Y}$.

If $A \in \mathscr{L}(\mathcal{G}, \mathcal{H})$, the **adjoint** or **conjugate**—at the risk of confusion with the dual or transpose A^*—is denoted by $A^\star \in \mathscr{L}(\mathcal{H}, \mathcal{G})$ and has the defining relation

$$\forall x \in \mathcal{H}\ \forall y \in \mathcal{G} : \quad \langle x|Ay\rangle_{\mathcal{H}} = \langle A^\star x|y\rangle_{\mathcal{G}}.$$

As a Hilbert space is a reflexive Banach space, the relations given for the transpose A^* in section 4.5 hold also for the adjoint $A^\star$.

Similarly to section 4.6, an operator $A \in \mathscr{L}(\mathcal{H})$ where $A^\star = A$ is called **self-adjoint** or **symmetric**, also **Hermitian** in the complex case. The subspace of self-adjoint operators is again denoted by $\mathrm{sym}(\mathcal{H}) \subset \mathscr{L}(\mathcal{H})$. If for $S \in \mathscr{L}(\mathcal{H})$ one has $S^\star = -S$, it is called **skew-adjoint**, the subspace of those operators again denoted by $\mathrm{so}(\mathcal{H}) \subset \mathscr{L}(\mathcal{H})$. Observe that on a complex space $\forall x \in \mathcal{H} : \langle x|Ax\rangle \in \mathbb{R}$; $\langle x|Sx\rangle \in \mathrm{i}\mathbb{R}$, whereas on a real space $\langle x|Sx\rangle = 0$. A self-adjoint operator $A \in \mathrm{sym}(\mathcal{H})$ is **positive (semi-definite)** if $\forall x \in \mathcal{H} : \langle x|Ax\rangle \geq 0$, and **positive definite** in case $\forall x \neq 0 : \langle x|Ax\rangle > 0$. Again, this defines a *partial order* on the subspace of self-adjoint operators, such that $M \leq A \Leftrightarrow A - M$ is positive. Observe that for any $A \in \mathscr{L}(\mathcal{H})$, the operators $A^\star A$ and $AA^\star$ are positive. Sometimes $|A| := \sqrt{A^\star A} = (A^\star A)^{1/2}$ is called the **absolute value** of A.

If for $U \in \mathscr{L}(\mathcal{H})$ it holds that $\forall x, y \in \mathcal{H} : \langle Ux|Uy\rangle = \langle x|y\rangle$, the operator U is **unitary**, or **orthogonal** in the real case. These operators form a group under composition, denoted by $\mathrm{U}(\mathcal{H})$, resp. $\mathrm{O}(\mathcal{H})$ in the real case. Observe that $U^{-1} = U^\star$, and that unitary and orthogonal operators are isometries. Remember the **polar decomposition** $A = U|A| = U(A^\star A)^{1/2}$, with U a partial isometry defined on $\mathrm{im}\,|A|$.

If $A \in \mathscr{L}(\mathcal{H})$ commutes with its adjoint $A^\star$, i.e. $[A, A^\star] = 0$ or $AA^\star = A^\star A$, it is called **normal**. Observe that self-adjoint, skew-adjoint, and unitary operators are normal. For a normal $A \in \mathscr{L}(\mathcal{H})$, the sub-algebra of $\mathscr{L}(\mathcal{H})$ generated by $\{A, A^\star\}$ is commutative. Note that a Hilbert space is reflexive, so the definitions of section 4.6 can be applied here as well. Hence there exists some potential for confusion, which can usually be clarified from the context. In case the Hilbert space is identified with its dual, all the notations agree with each other.

Remember that A is compact ($A \in \mathscr{L}_0(\mathcal{H})$) iff there are two orthonormal systems (ONS) $\{e_j\}_{j\in J}$ and $\{f_j\}_{j\in J}$, i.e. $\forall j,k \in J : \langle e_j|e_k\rangle = \delta_{jk}$, and a sequence of numbers $(\varsigma_j)_{j\in J} \in \boldsymbol{c}_0$ —see section 6.1—the **singular values**—see section 4.8—such that with a series converging in $\mathcal{H}$

$$\forall x \in \mathcal{H} : \; Ax = \sum_{j\in J} \varsigma_j \langle e_j|x\rangle f_j.$$

This is called the **singular value decomposition** (SVD), or sometimes the Schmidt decomposition. (For a complete orthonormal system (CONS) in (a separable) $\mathcal{H}$ one has additionally that $\overline{\operatorname{span}}\{e_j\}_{j\in J} = \mathcal{H}$.) The operator norm of A is then $\|A\|_{\mathscr{L}} = \|(\varsigma_j)\|_\infty$. In the following context, the operator norm $\|A\|_{\mathscr{L}}$ may also be denoted by $\|A\|_{(\infty)}$. If only finitely many ς_j are non-zero ($(\varsigma_j) \in \boldsymbol{c}_{00}$), A is a finite rank operator —$\dim(\operatorname{im} A) \in \mathbb{N}_0$—$A \in \mathscr{L}_{00}(\mathcal{H})$.

If in addition $\|(\varsigma_j)\|_2 < \infty$ (i.e. $(\varsigma_j) \in \ell_2$, see section 6.1), the operator $A \in \mathscr{L}_2(\mathcal{H})$is called a **Hilbert-Schmidt** operator. If $\|(\varsigma_j)\|_1 < \infty$, i.e. $(\varsigma_j) \in \ell_1$, then $A \in \mathscr{L}_1(\mathcal{H})$ is of **trace-class** or **nuclear**, with **trace** $\operatorname{tr} A := \sum_{j\in J}\langle Ae_j|e_j\rangle$. More generally, the **Schatten**-classes $\mathscr{L}_p(\mathcal{H})$ may be defined (where $1 \le p < \infty$) as those $A \in \mathscr{L}_0(\mathcal{H})$ such that $\|A\|_{(p)} := \|(\varsigma_j)\|_p < \infty$, i.e. $(\varsigma_j) \in \ell_p$.

Obviously for $1 \le p \le q < \infty$ one has $\mathscr{L}_{00}(\mathcal{H}) \subseteq \mathscr{L}_p(\mathcal{H}) \subseteq \mathscr{L}_q(\mathcal{H}) \subseteq \mathscr{L}_0(\mathcal{H}) \subseteq \mathscr{L}(\mathcal{H})$ with continuous embeddings, and especially $\mathscr{L}_1(\mathcal{H}) \subseteq \mathscr{L}_2(\mathcal{H})$. Remember that for the natural duality pairing $\langle A, B\rangle_{\mathscr{L}} := \operatorname{tr}(B^*A)$ one has $\mathscr{L}_0(\mathcal{H})^* \cong \mathscr{L}_1(\mathcal{H})$, $\mathscr{L}_1(\mathcal{H})^* \cong \mathscr{L}(\mathcal{H})$, and $\mathscr{L}_2(\mathcal{H})^* \cong \mathscr{L}_2(\mathcal{H})$ is a Hilbert space. More generally, for $1 < p < \infty$ and the **dual exponent** p^* such that $1/p + 1/p^* = 1$, one has $\mathscr{L}_p(\mathcal{H})^* \cong \mathscr{L}_{p^*}(\mathcal{H})$, showing that those $\mathscr{L}_p(\mathcal{H})$ are reflexive spaces, see section 4.6.

4.7 Unbounded Linear Operators

A partial (see section 1.6) map A from a Hilbert space $\mathcal{H}$ into itself is called an operator *in* a Hilbert space, as opposed to an operator *on* a Hilbert space (like those in $\mathscr{L}(\mathcal{H})$. If the **domain of definition** $\operatorname{dom} A$ is a *subspace* of the Hilbert space $\mathcal{H}$ and A is linear on $\operatorname{dom} A$, then A is called a *linear operator in* $\mathcal{H}$. Observe that in this case

1. the **graph** of A (note the difference to section 4.4) $\operatorname{gra} A = \{(x, y) : x \in \operatorname{dom} A,\ y = Ax\}$ is a subspace of $\mathcal{H} \times \mathcal{H} \cong \mathcal{H} \oplus \mathcal{H}$,

2. and $(0, x) \in \operatorname{gra} A \Leftrightarrow x = 0$.

Conversely, if a subspace of $\mathcal{H} \oplus \mathcal{H}$ satisfies the second condition, it defines a linear operator in $\mathcal{H}$. In case $\mathrm{cl}(\mathrm{dom}\, A) = \mathcal{H}$, i.e. the domain of A is dense in $\mathcal{H}$, one says that A is **densely defined.** Observe that if a densely defined A is bounded, it has a unique continuous extension to all of $\mathcal{H}$. Hence the interesting case are *unbounded* operators.

In the following, the direct sum $\mathcal{H} \oplus \mathcal{H}$ is equipped with the appropriate inner product, see section 7. On the subspace $\mathrm{gra}\, A \subseteq \mathcal{H} \oplus \mathcal{H}$ this is called the *graph* inner product $\langle x|y\rangle_{\mathcal{H}} + \langle Ax|Ay\rangle_{\mathcal{H}}$. The projection on the first component, which is dom A, is usually equipped with the **energy** or *operator inner product* $\langle x|y\rangle_A := \langle Ax|Ay\rangle_{\mathcal{H}}$. An operator A is

closed iff $\mathrm{gra}\, A$ is a closed subspace of $\mathcal{H} \oplus \mathcal{H}$ in the graph inner product.

isometric or an *isometry* iff $\forall x, y \in \mathrm{dom}\ A : \langle Ax|Ay\rangle = \langle x|y\rangle$.

symmetric iff $\forall x, y \in \mathrm{dom}\ A : \langle Ax|y\rangle = \langle x|Ay\rangle$.

An operator A' is called an **extension** of A —denoted by $A \subset A'$ —iff gra $A \subset$ gra A'. If A admits a closed extension, it is called **closable.** If cl(gra A) defines an operator (the second condition above), this is called the **closure** $\bar{A}$ of A, and is the smallest closed extension.

Denote the unitary mapping $(x, y) \mapsto (-y, x)$ on $\mathcal{H} \oplus \mathcal{H}$ by J, then for a densely defined operator A define the **adjoint** as the operator A^* with gra $A^\star = (J(\mathrm{gra}\ A))^\perp = J((\mathrm{gra}\ A)^\perp)$. Note that $A^\star$ is closed. In case A is closed, $A^\star$ is also densely defined, so that $A^{\star\star}$ exists, and $A^{\star\star} = \bar{A}$.

If A is densely defined and symmetric, $A \subset A^\star$. If $A^\star = \bar{A}$, the operator A is called **essentially self-adjoint.** In case also dom $A =$ dom $A^\star$, the operator A is called **self-adjoint**, or also **Hermitian** in the complex case. Note that in this case A is necessarily closed, as gra $A =$ gra $A^\star$. In the essentially self-adjoint case, observe that $A^\star$ is self-adjoint. For a densely defined operator A, in case only dom $A =$ dom $A^\star$ and $AA^\star = A^\star A$ hold, A is called **normal**, and satisfies $\|Ax\| = \|A^* x\|$ for $x \in \mathrm{dom}\ A$.

A symmetric linear operator A is called **semi-bounded below**, iff $\exists c \in \mathbb{R}\ \forall x \in \mathrm{dom}\ A :\ \langle Ax|x\rangle \geq c\|x\|^2$. Observe that positive operators are semi-bounded below with $c = 0$. If $c > 0$, the operator is **coercive** and hence positive definite, it has a bounded inverse. Note that a densely defined symmetric and semi-bounded operator admits a self-adjoint extension (Friedrichs' extension).

For a densely defined symmetric coercive—hence self-adjoint—operator A in $\mathcal{H}$, there are unique positive self-adjoint operators A^s for $s \in \mathbb{R}_*$ in $\mathcal{H}$ (the positive s-th power). Set $\mathcal{G} := \bigcap_{s \geq 0}(\mathrm{dom}\ A^s)$, define the 'operator norms' $\|u\|_s := \sqrt{\langle A^s u|u\rangle_{\mathcal{H}}}$ and $\mathcal{H}_s = \mathrm{cl}_s \mathcal{G}$. Then $\{\mathcal{H}_s\}_{s \in \mathbb{R}_*}$ is a **scale** of Hilbert spaces with norms $\|\cdot\|_s$, such that

$$\mathcal{H}_s \hookrightarrow \mathcal{H}_t \hookrightarrow \mathcal{H} \cong \mathcal{H}_0 \text{ for } s > t > 0,$$

with dense embeddings. Identify $\mathcal{H}$ with its dual $\mathcal{H}^*$ —the **pivot space**. Denote the dual of $\mathcal{H}_s$ by $\mathcal{H}_{-s}$, then $\mathcal{H} \hookrightarrow \mathcal{H}_{-s}$ densely, and one has a **Gelfand triplet** or **rigged Hilbert space**

$$\mathcal{H}_s \hookrightarrow \mathcal{H} \cong \mathcal{H}_0 \hookrightarrow \mathcal{H}_{-s},$$

or more generally for $s > t > 0$:

$$\mathcal{G} \hookrightarrow \mathcal{H}_s \hookrightarrow \mathcal{H}_t \hookrightarrow \mathcal{H} \hookrightarrow \mathcal{H}_{-t} \hookrightarrow \mathcal{H}_{-s} \hookrightarrow \mathcal{G}^*.$$

Each A^s may now be extended to a *continuous* linear map $A^s : \mathcal{H}_s \to \mathcal{H}_{-s}$, in particular $A \in \mathscr{L}(\mathcal{H}_1, \mathcal{H}_{-1})$. The duality pairing for $x \in \mathcal{H}_s$, $y \in \mathcal{H}$ may be computed via the inner product of the pivot space $\langle x, y\rangle_{\mathcal{H}_s \times \mathcal{H}_{-s}} = \langle x|y\rangle_{\mathcal{H}}$. Observe that $\mathcal{G}$ together with the norms $\|\cdot\|_{A^k}$, $k \in \mathbb{N}_0$ is a countably Hilbert space.

4.8 Spectrum and Singular Values

The **resolvent** of a linear operator A is the operator valued map $z \mapsto R_A(z) := (zI - A)^{-1}$ for $z \in \mathbb{C}$. The **resolvent set** is the subset of $\mathbb{C}$ where $R_A(z)$ is everywhere defined and continuous, hence $\varrho(A) = \{z \in \mathbb{C} : R_A(z) \text{ exists in } \mathscr{L}(\mathcal{H})\}$. The complement $\sigma(A) := \complement\varrho(A) = \mathbb{C} \setminus \varrho(A)$ is the **spectrum** of A, where $R_A(z)$ either does not exist, or is not continuous or not defined on all of $\mathcal{H}$. The spectrum is a closed set, for a bounded operator it is also bounded, hence compact, and $r(A) := \max\{|z| : \ z \in \sigma(A)\}$ is the **spectral radius.**

The **point spectrum** $\sigma_p(A) \subseteq \sigma(A)$ are those z where $(zI - A)$ is not injective—z is an eigenvalue, usually denoted by λ. The **continuous spectrum** $\sigma_c(A)$ is where $(zI - A)$ is injective, but not surjective and $\operatorname{im}(zI - A)$ is dense in $\mathcal{H}$, so $R_A(z)$ may be seen as a densely defined unbounded operator. The **residual spectrum** $\sigma_r(A)$ is the rest where $(zI - A)$ is injective, but not surjective and $\operatorname{im}(zI - A)$ is not dense in $\mathcal{H}$, so that $\sigma(A) = \sigma_p(A) \cup \sigma_c(A) \cup \sigma_r(A)$.

The **approximate point spectrum** $\sigma_{ap}(A)$ are those $z \in \sigma(A)$, where for a *Weyl* sequence $x_n \in \mathcal{H}$ with $\|x_n\| = 1$ one has $\|(zI - A)x_n\| \to 0$, i.e. $(zI - A)$ has no continuous inverse. Thus $\sigma_p(A), \sigma_c(A) \subseteq \sigma_{ap}(A) \subseteq \sigma(A)$. The **absolutely continuous spectrum** $\sigma_{ac}(A)$ is where the spectral measure $\mathsf{s}(dz) = \langle E_{dz}x|x\rangle$ is absolutely continuous w.r.t. Lebesgue measure dz. The **peripheral spectrum** $\sigma_\pi(A)$ are those $z \in \sigma(A)$ with $|z| = r(A)$.

The **discrete spectrum** $\sigma_d(A) \subseteq \sigma_p(A)$ are the isolated points $z \in \sigma(A)$ where $(zI - A) \in \mathscr{L}_\Phi(\mathcal{H})$ is a Fredholm operator. The **essential spectrum** is the rest $\sigma_e(A) = \sigma(A) \setminus \sigma_d(A)$.

The **singular values** of an operator $A \in \mathscr{L}_0(\mathcal{H})$ are the positive square roots of the eigenvalues of $A^\star A \in \mathscr{L}_0(\mathcal{H})$, and could be denoted by

$$\varsigma(A) := \{s \in \mathbb{R}_+ : s^2 \in \sigma_p(A^\star A)\}.$$

5 Ordered Vector Spaces

A convex and salient cone $\mathcal{C}$ defines a partial order on $\mathcal{X}$ via $x \preceq y \Leftrightarrow y - x \in \mathcal{C}$. In this context, $\mathcal{C}$ is the **positive cone** of elements $0 \preceq x$, sometimes denoted by $\mathcal{X}_+$.

The order is compatible with the linear structure, namely for all $x \in \mathcal{X}$ the relation $y \preceq z$ implies $x + y \preceq x + z$, and from $0 \preceq x$ follows that for all $\alpha \in \mathbb{R}_+$ also $0 \preceq \alpha x$.

A vector space with a compatible partial order which is also a lattice—see section 1.5—is called a **Riesz space**. Here one may define for each $x \in \mathcal{X}$ the **positive part** $x^+ := x \vee 0$, the **negative part** $x^- := (-x) \vee 0$ (observe that $0 \preceq x^-$), and the **absolute value** (really a vector, not a number) $\mathcal{X} \ni |x| := x \vee (-x)$. Observe that $x = x^+ - x^-$ and $|x| = x^+ + x^-$, as well as $x + y = x \vee y + x \wedge y$, and $|x - y| = x \vee y - x \wedge y$, and $x^+ \wedge x^- = 0$. Two vectors which satisfy $|x| \vee |y| = 0$ are sometimes termed **disjoint** or **order orthogonal**, denoted by $x \perp y$. The infimum should not be confused with the *wedge* or *exterior product* (see section 8.3).

A norm or seminorm on $\mathcal{X}$ such that $|x| \preceq |y|$ imples $\|x\| \leq \|y\|$ is called a *Riesz* norm or seminorm. A Riesz space with a Riesz norm is a **normed Riesz space**. If it is complete w.r.t. the norm, it is called a **Banach lattice**.

A net $\{x_j\}_{j \in J} \subset \mathcal{X}$ is **decreasing**, written $x_j \downarrow$, if $x_k \preceq x_j$ whenever $j \preceq k$. Similarly for an **increasing** net $x_j \uparrow$. The notation $x_j \downarrow x$ means that $x = \inf\{x_j\}$, and similarly $x_j \uparrow x$ means that $x = \sup\{x_j\}$. The **limes inferior** is $\liminf_{j \in J} x_j := \sup_{j \in J} \inf_{j \preceq k} x_k$, and analogous for the **limes superior** $\limsup_{j \in J} x_j := \inf_{j \in J} \sup_{j \preceq k} x_k$. The net x_j is **order convergent** to $x \in \mathcal{X}$, denoted by $x_j \xrightarrow{o} x$, if there is a net $\{y_j\}_{j \in J}$ on the same directed set J with $y_j \downarrow 0$ such that $|x - x_j| \preceq y_j$.

Generalised intervals or rectangles in ordered vector spaces are defined via $[x, y] := \{z : x \preceq z \wedge z \preceq y\}$, or as $]x, y] = (x, y] := \{z : x \prec z \wedge z \preceq y\}$, and similarly other combinations. The notation $[x, y]$ could be confused with the line segment in section 4.5, but it should be clear from the context what is meant.

An isotone (see section 1.5) linear functional x^* is **positive**, as $0 \preceq x$ implies $0 = \langle x^*, 0\rangle \leq \langle x^*, x\rangle$. If $\mathcal{C} = \mathcal{X}_+$ is the order defining positive cone in $\mathcal{X}$, the positive polar cone $\mathcal{C}^\oplus$ (see section 4.5) is likewise a closed, convex, and salient cone in $\mathcal{X}^*$, defining that order relation in $\mathcal{X}^*$, i.e. $\mathcal{X}^*_+ = \mathcal{C}^\oplus$.

If for some $0 \preceq y$ the relation $x \in [-\alpha y, \alpha y]$ for all $\alpha \in \mathbb{R}_*$ implies $x = 0$, the Riesz space is called **almost Archimedean**. If in the same situation the stronger relation $x \in [0, \alpha y]$ for all $\alpha \in \mathbb{R}_*$ implies $x = 0$, the Riesz space is called **Archimedean**.

An element $e \in \mathcal{C}$ ($0 \preceq e$) is called an **order unit** for $\mathcal{C}$ if for each $x \in \mathcal{X}$ there is an $\alpha \in \mathbb{R}_*$ such that $x \in [-\alpha e, \alpha e]$. If $\mathcal{X}$ has an order unit, then

$$\|x\|_o := \inf\{\alpha \in \mathbb{R}_* : \ x \in [-\alpha e, \alpha e]\}$$

defines an **order unit norm** exactly when $\mathcal{X}$ is almost Archmedean. If $\mathcal{X}$ is actually Archimedean, then the closed unit ball is $[e, e]$.

Any $\phi \in \mathcal{X}^*$ (the dual of $(\mathcal{X}, \|\cdot\|_o)$) has a decomposition $\phi = \phi^+ - \phi^-$ with $\phi^+, \phi^- \in \mathcal{C}^\oplus$ such that $\|\phi\|_o = \|\phi^+\|_o + \|\phi^-\|_o$.

If $\mathscr{G}(\mathcal{X}, \mathcal{Y}) \subseteq \mathscr{F}(\mathcal{X}, \mathcal{Y})$ is some space of functions into an ordered vector space $\mathcal{Y}$ (often $\mathcal{Y} = \mathbb{R}$), then one may define an order relation on $\mathscr{G}(\mathcal{X}, \mathcal{Y})$ via $f \leq g \ \Leftrightarrow \ \forall x \in \mathcal{X} : \ f(x) \preceq g(x)$. The positive cone are those functions f where $0 \preceq x$ implies $0 \preceq f(x)$.

6 Examples of Vector Spaces

Some examples of vector spaces were already noted, namely $\mathbb{K}^n$, $\boldsymbol{c}_{00}$, $\boldsymbol{\omega}$, $\mathbb{K}^{n\times n}$, and the space $\mathscr{L}(\mathcal{X}, \mathcal{Y})$, or $\mathscr{L}(\mathcal{H})$ and its subspaces from section 4. In case $\mathcal{Y}$ is a (topological) vector space and $\mathcal{X}$ any set, the spaces $\mathscr{F}(\mathcal{X}, \mathcal{Y})$, $\mathscr{B}(\mathcal{X}, \mathcal{Y})$, and $C(\mathcal{X}, \mathcal{Y})$ —for a topological space $\mathcal{X}$ —and its subspaces from section 2 are (topological) vector spaces—as subspaces of $\mathscr{F}(\mathcal{X}, \mathcal{Y}) = \mathcal{Y}^{\mathcal{X}}$, see section 4.2. More important topologies arise from introducing appropriate *seminorms* and *norms* (see section 4.1), which make these spaces *complete*.

6.1 Sequence Spaces

An important example is $\mathbb{R}^n \subset \mathbb{R}^{(\mathbb{N})} = \boldsymbol{c}_{00} \subset \mathbb{R}^{\mathbb{N}} = \boldsymbol{\omega}$. This last space is usually equipped with the projective product topology, whereas $\boldsymbol{c}_{00} = \mathbb{R}^{(\mathbb{N})}$ then carries the injective topology. The natural duality pairing between $x = (\xi_n)$ and $y = (\eta_n)$ from some sequence space is $\langle x, y\rangle := \sum_{n\in\mathbb{N}} \xi_n \eta_n$. Hence $\boldsymbol{\omega}^* \cong \boldsymbol{c}_{00}$.

Other subspaces are $\ell_\infty := \{(\varrho_n) : \|(\varrho_n)\|_\infty := \sup_n |\varrho_n| < \infty\} \subset \boldsymbol{\omega}$, the space of **bounded sequences**, as well as in ℓ_∞ the closed subspace of

convergent sequences $\boldsymbol{c} := \{(\varrho_n) : \ \exists \rho := \lim_n \varrho_n\} \subset \ell_\infty$, and its closed subspace of **zero-convergent sequences** $\boldsymbol{c}_0 := \{(\varrho_n) : \ \rho := \lim_n \varrho_n = 0\} \subset \boldsymbol{c}$. Observe $\boldsymbol{c} = \mathbb{R} \oplus \boldsymbol{c}_0$ with $\boldsymbol{c} \ni (\varrho_n) \mapsto (\rho, (\varrho_n - \rho)) \in \mathbb{R} \oplus \boldsymbol{c}_0$.

Very similar to ℓ_∞ and $\boldsymbol{c}$ is the space of **bounded series**

$$\boldsymbol{bs} := \left\{ (\varrho) \in \ell_\infty | \ \sup_n \left| \sum_{j=1}^n \varrho_j \right| < \infty \right\}.$$

With the map $\boldsymbol{bs} \ni (\varrho_n) \mapsto (\sum_{j=1}^n \varrho_j) \in \ell_\infty$ this space is isometrically isomorphic to ℓ_∞. It has the closed subspace of **convergent series**

$$\boldsymbol{cs} := \left\{ (\varrho) \in \ell_\infty | \ \exists \varsigma := \lim_n \left| \sum_{j=1}^n \varrho_j \right| < \infty \right\},$$

which is mapped onto $\boldsymbol{c}$ under above mapping. Another subspace is

$$\boldsymbol{s} := \{(\varrho_n) | \ \forall k \in \mathbb{N} : \ (n^k \varrho_n) \in \ell_1\},$$

the space of **rapidly decreasing sequences**, equipped with the seminorms $|(\varrho_n)|_k := \sum_n |n^k \varrho_n|$, and its dual of **tempered sequences**

$$\boldsymbol{s}^* := \{(\varrho_n) | \ \exists k \in \mathbb{N} : \ (n^{-k} \varrho_n) \in \ell_\infty\}.$$

Other well known examples are the ℓ_p-spaces for $0 < p < \infty$

$$\ell_p := \left\{ (\varrho_n) \in \mathbb{R}^{\mathbb{N}} : \ \|(\varrho_n)\|_p^p := \sum_{n=1}^\infty |\rho_n|^p < \infty \right\}.$$

For $0 < p < 1$ these are (metrisable) topological vector spaces (TVS)—as $\| \cdot \|_p^p$ is a F-norm, but not locally convex, whereas for $1 \le p \le \infty$ they are Banach spaces. For $1 \le p < q \le \infty$ the embeddings $\ell_p \to \ell_q$ are compact.

One may set $\ell_{0+} := \bigcap_{p>0} \ell_p$. The product space $\boldsymbol{\omega}$ is not normable, but **metrisable** with either of the translation invariant metrics or F-norms (compare section 3 with the metric on L_0).

$$d((\varrho_n), (\rho_n)) := \sum_n \frac{|\varrho_n - \rho_n|}{2^n (1 + |\varrho_n - \rho_n|)},$$

or

$$d((\varrho_n), (\rho_n)) := \sum_n \frac{\max\{|\varrho_n - \rho_n|, 1\}}{2^n}.$$

For $0 < p < q < \infty$: $\boldsymbol{c}_{00} \hookrightarrow \boldsymbol{s} \hookrightarrow \ell_p \hookrightarrow \ell_q \hookrightarrow \boldsymbol{c}_0 \hookrightarrow \boldsymbol{c} \hookrightarrow \ell_\infty \hookrightarrow \boldsymbol{s}^* \hookrightarrow \boldsymbol{\omega}$. For $p > 1$ and p^* satisfying $1/p + 1/p^* = 1$, $\ell_p^* \cong \ell_{p^*}$. If $\mathbb{R}^n$ is equipped with an ℓ_p-norm, the space is often denoted by $\ell_p^n \hookrightarrow \ell_p$. Obviously the spaces ℓ_p and its subspaces ℓ_p^n for $1 < p < \infty$ are *reflexive*, see section 4.6, and ℓ_2 is a Hilbert space.

Also remember that in analogy to section 4, $\boldsymbol{c}_0^* \cong \ell_1$, $\boldsymbol{c}^* \cong (\mathbb{R} \oplus \boldsymbol{c}_0)^* \cong \mathbb{R} \times \ell_1$, $\ell_1^* \cong \ell_\infty$, and . As already stated, $\boldsymbol{c}_{00} \cong \boldsymbol{\omega}^*$. The dual of ℓ_∞ is ismophic to $\boldsymbol{ba} \cong \ell_\infty^*$, the space of bounded additive measures on $2^{\mathbb{N}}$.

6.2 Spaces of Continuous and Differentiable Maps

The product topology inherited from $\mathcal{Y}^{\mathcal{X}}$ is the *simple topology* of *pointwise convergence*. Pointwise convergence of a sequence of functions $\{f_n\} \in \mathscr{B}(\mathcal{X})$ corresponds to weak convergence, similarly for the closed subspace of measurable bounded functions $\mathscr{B}_b(\mathcal{X})$. But $C_b(\mathcal{X})$ is not a closed subspace with this topology, and neither is $C(\mathcal{X}, \mathcal{Y})$.

The space $\mathscr{B}(\mathcal{X})$ of bounded functions is a Banach space with the norm $\|f\|_\infty = \sup_{x \in \mathcal{X}} |f(x)|$, as is its closed subspace $\mathscr{B}_b(\mathcal{X})$ of bounded measurable functions. Its dual space is $\boldsymbol{ba}(\mathcal{X}, \mathfrak{A}) = \mathscr{B}_b(\mathcal{X})^*$, the space of signed charges on the algebra $\mathfrak{A}$ generated by the open sets. The space $C_b(\mathcal{X})$ is a closed subspace of $\mathscr{B}_b(\mathcal{X})$, and so is $C_0(\mathcal{X})$. Of course (see section 5) $\mathscr{B}(\mathcal{X})$, $\mathscr{B}_b(\mathcal{X})$, $C_b(\mathcal{X}), C_0(\mathcal{X})$, and $C_{00}(\mathcal{X})$ are also Riesz spaces, and pointwise convergence also corresponds to order convergence.

If $\mathcal{X}$ is a totally bounded metric space, and $0 < \alpha < \beta \leq 1$, then for the Hölder spaces $C^{0,\alpha}(\mathcal{X})$, which are Banach spaces (see section 2.2) with the norm $|\cdot|_\alpha$, we have the following continuous, dense, and compact embedding:

$$C^{0,\beta}(\mathcal{X}) \hookrightarrow C^{0,\alpha}(\mathcal{X}) \hookrightarrow C_b(\mathcal{X}).$$

For $\mathcal{K} \subseteq \mathcal{X}$, and $\mathcal{K}$ compact, define the seminorm $|\varphi|_{C(\mathcal{K})} := \sup_{x \in \mathcal{K}} |\varphi(x)|$. The compact open topology is then provided by the *family of seminorms* $\{|\varphi|_{C(\mathcal{K})} : \mathcal{K} \subseteq \mathcal{X} \text{ compact}\}$, the topology of uniform convergence on compact sets. The space is complete with this topology. The same is true for the closed subspaces $C_u(\mathcal{X})$ and $C_{00}(\mathcal{X})$. Unless $\mathcal{X}$ is compact (where the seminorm with $\mathcal{K} = \mathcal{X}$ is already a norm and sufficient to define the topology), this topology is *not metrisable*.

Similarly, if $\mathcal{X}$ is a differentiable manifold, the spaces of k-times differentiable maps are denoted by $C^k(\mathcal{X})$, and are usually equipped with the family of seminorms $\{|\varphi|_{C^k(\mathcal{K})} := \max_{0 \leq |\mathfrak{k}| \leq k} |\mathrm{D}^{\mathfrak{k}} \varphi|_{C(\mathcal{K})} : \mathcal{K} \subseteq \mathcal{X} \text{ compact}\}$. The same applies to the space

$$C_u^k(\mathcal{X}) := \{\phi |\ \phi \in C^k(\mathcal{X}) \wedge \mathrm{D}^j \phi \in C_u(\mathcal{X}), j = 0, \ldots, k\},$$

and the completely analogously defined space $C^k_{00}(\mathcal{X})$. The similarly defined subspaces $C^k_b(\mathcal{X})$ and $C^k_0(\mathcal{X})$ of $C^k(\mathcal{X})$ are non-reflexive *Banach spaces* with the norm $\|\varphi\|_{C^k} := \sup_{0\le|\mathfrak{k}|\le k}\|\mathrm{D}^{\mathfrak{k}}\varphi\|_\infty$. If $\mathcal{X}$ is compact, then for any $k\in\mathbb{N}$: $C^k(X) = C^k_u(X) = C^k_b(\mathcal{X}) = C^k_0(\mathcal{X})$. The space $C^k_{00}(\mathcal{X})$ is in this case often defined as those functions in $C^k(\mathcal{X})$ whose support is a proper subset of $\mathcal{X}$.

For any of the Banach spaces just described, e.g. $C^k_b(\mathcal{X})$, define

$$C^{k,\beta}_b(\mathcal{X}) := \{\phi\in C^k_b(\mathcal{X}):\ |\mathrm{D}^k\phi|_\beta<\infty\},$$

here with the norm $\|\phi\|_{C^{k,\beta}} := \max\{\|\phi\|_{C^k}, |\mathrm{D}^k\phi|_\beta\}$. Again for $0<\alpha<\beta\le 1$ and any $k\in\mathbb{N}$, we have the continuous, dense, and compact embedding:

$$C^{k,\beta}_b(\mathcal{X}) \hookrightarrow C^{k,\alpha}_b(\mathcal{X}) \hookrightarrow C^k_b(\mathcal{X}).$$

Also observe $\mathscr{E}(\mathcal{X}) := C^\infty(\mathcal{X}) := \bigcap_{k\ge0} C^k(\mathcal{X})$ with the *projective limit topology* given by the family of seminorms $\{|\varphi|_{C^k(\mathcal{K})}:\ k\in\mathbb{N}_0, \mathcal{K}\subseteq\mathcal{X} \text{ compact }\}$ —this space is *not metrisable* (see section 4.1) unless $\mathcal{X}$ is compact—and similarly for $\mathscr{D}(\mathcal{X}) := C^\infty_{00}(\mathcal{X}) := \bigcap_{k\ge0} C^k_{00}(\mathcal{X})$ with corresponding seminorms. In case $\mathcal{X}$ is pre-compact, the limit spaces are *nuclear spaces*, albeit with different families of norms. Observe that they are all associative Abelian algebras under point-wise multiplication of functions.

Their duals are—a bit inconsistently here but conventionally—denoted by $\mathscr{E}'(\mathcal{X}) := \mathscr{E}(\mathcal{X})^*$ for the space of compactly supported distributions, and $\mathscr{D}'(\mathcal{X}) := \mathscr{D}(\mathcal{X})^*$ for the usual space of distributions. They both carry the *injective limit topology*.

If the set $\mathcal{X}$ is a domain in $\mathbb{R}^n$, the *Schwartz* space of **rapidly decreasing** smooth functions is $\mathscr{S}(\mathcal{X})\subset\mathscr{E}(\mathcal{X})$ with the countable family of semi-norms $|\varphi|_{\mathscr{S},k,\ell} := \sup\{\|x^j\mathrm{D}^m\varphi\|_\infty:\ j,m\in\mathcal{N}_0,\ |m|\le\ell,\ |j|\le k\}$ —a *Fréchet* and also a *nuclear space*, and its dual $\mathscr{S}'(\mathcal{X}) := \mathscr{S}(\mathcal{X})^*$ is the *Schwartz* space of **tempered distributions**. Note the continuous embeddings $\mathscr{D}(\mathcal{X})\hookrightarrow\mathscr{S}(\mathcal{X})\hookrightarrow\mathscr{E}(\mathcal{X})$ and $\mathscr{E}'(\mathcal{X})\hookrightarrow\mathscr{S}'(\mathcal{X})\hookrightarrow\mathscr{D}'(\mathcal{X})$.

The natural duality pairing, say between $\varphi\in\mathscr{S}'(\mathcal{X})$ and $f\in\mathscr{S}(\mathcal{X})$ is usually written formally

$$\int_{\mathcal{X}}\varphi f\,\mathrm{d}\lambda = \int_{\mathcal{X}}\varphi(x)f(x)\,\mathrm{d}x := \langle\varphi,f\rangle,$$

where λ is Lebesgue measure on $\mathcal{X}$.

6.3 Lebesgue Spaces

For a measure space $(\Omega,\mathfrak{F},\mu)$ for $0<p<\infty$ the **Lebesgue** spaces are

$$L_p(\Omega,\mathfrak{F},\mu) := \left\{ f\in L_0(\Omega,\mathfrak{F}) : \|f\|_p^p := \int_\Omega |f(\omega)|^p\,\mu(\mathrm{d}\omega)<\infty\right\},$$

and $L_\infty(\Omega, \mathfrak{F}, \mu) := \{f \in L_0(\Omega, \mathfrak{F}) : \operatorname{ess\,sup}_{\omega\in\Omega} |f(\omega)| < \infty\}$, which is also an associative Abelian algebra under point-wise multiplication. Obviously $\mathscr{B}_b(\Omega, \mathfrak{F})$ and its subspaces $C_b(\Omega)$ and $C_c(\Omega)$ are *closed* subspaces of $L_\infty(\Omega, \mathfrak{F}, \mu)$. In case the σ-algebra $\mathfrak{F}$ is clear, it may be omitted, and analogous also the measure μ or the base space Ω, i.e. $L_p(\Omega, \mathfrak{F}, \mu) = L_p(\Omega, \mu) = L_p(\Omega) = L_p(\mu)$.

For $0 < p < 1$ these are (metrisable) TVS—as $\|\cdot\|_p^p$ is a F-norm, but not locally convex, whereas for $1 \le p \le \infty$ they are Banach spaces. Remember that for $0 < p < q \le \infty$ and a finite measure space ($\mu(\Omega) < \infty$), one has the continuous embeddings $L_q(\Omega) \hookrightarrow L_p(\Omega)$. The natural duality pairing for $f \in L_p(\Omega)$ and $g \in L_q(\Omega)$ is $\int_\Omega f\, g \,\mathrm{d}\mu = \int_\Omega f(\omega) g(\omega)\, \mu(\mathrm{d}\omega)$.

And for $1 < p < \infty$ and p^* as before, $L_p(\Omega)^* \cong L_{p^*}(\Omega)$ for any measure space, which means that these spaces are *reflexive*, see section 4.6. Also remember that $L_1(\Omega)^* \cong L_\infty(\Omega)$ for a σ-finite measure space. $L_2(\Omega)$ is a Hilbert space. Note also $L_{q+}(\Omega) := \bigcup_{p>q} L_p(\Omega)$ and $L_{q-}(\Omega) := \bigcap_{p<q} L_p(\Omega)$.

If in addition Ω is a locally compact group with **Haar** measure λ, define the **convolution** $f * g$ of $f, g \in L_1(\Omega)$ as $(f * g)(y) := \int_\Omega f(x) g(y - x)\, \lambda(\mathrm{d}x)$, and the **correlation** as $(f \star g)(y) := \int_\Omega f(x) g(y + x)\, \lambda(\mathrm{d}x)$. The space $L_1(\Omega)$ together with the convolution is an associative **Abelian** algebra.

The spaces $L_p(\Omega)$ are Riesz spaces under the pointwise ordering in $\mathbb{R}$, i.e. $f \le g \Leftrightarrow f(x) \le g(x)\ \forall x \in \Omega \setminus N$, where N is a null-set ($\mu(N) = 0$). The convergence **almost everywhere** (abbreviated a.e. and in probability theory called **almost surely**, abbreviated a.s.), which means that $f_n(\omega) \to f(\omega)$ a.e. $\Leftrightarrow f_n(\omega) \to f(\omega)\, \forall \omega \in \Omega \setminus N,\ \mu(N) = 0$, is convergence in the metric topology inherited as a subset of $L_0(\Omega)$ —see section 3, and is the same as the order convergence in the Riesz space, see section 5.

6.4 Sobolev Spaces

If Ω is also a differentiable manifold, the corresponding **Sobolev** spaces with up to k-th distributional derivative in $L_p(\Omega)$ are denoted by $W_p^k(\Omega) \hookrightarrow L_p(\Omega)$ with norm for $u \in W_p^k(\Omega)$

$$\|u\|_{W_p^k} := \left(\sum_{j=0}^{k} \|\mathrm{D}^j u\|_{L_p}^p \right)^{1/p}.$$

For $1 < p < \infty$ these spaces are *reflexive*, whereas for $p = 2$ these are Hilbert spaces, with a special notation $H^k(\Omega) := W_2^k(\Omega)$.

The completion of $C_{00}^\infty(\Omega)$ in the W_p^k-norm is denoted by $\mathring{W}_p^k(\Omega)$, and in the case of $p = 2$ by $\mathring{H}^k(\Omega) := \mathring{W}_2^k(\Omega)$. The duals of these last spaces are denoted by $\mathring{W}_p^k(\Omega)^* \cong W_{p^*}^{-k}(\Omega)$ and $\mathring{H}^k(\Omega)^* \cong H^{-k}(\Omega)$ respectively, where the natural duality pairing is the same as in the case of Lebesgue spaces in section 6.3. On these spaces, $|u|_{k,p} := \|\mathrm{D}^k u\|_{L_p}$ is an equivalent norm, whereas on the whole space it is only a semi-norm.

Sobolev spaces may be defined for non-integer orders, and simplest is to start with the Hilbert space $H^1(\Omega)$. Following section 4.7 we see that in $\mathcal{H} = L_2(\Omega)$ the H^1-inner product defines a symmetric positive definite bilinear form and therefore an unbounded self-adjoint positive definite operator

$$\langle u|v\rangle_{H^1} = \langle Bu|v\rangle_{\mathcal{H}}.$$

It hence has a unique square root $A = B^{1/2}$. If now as in section 4.7 we take A as the defining operator, we obtain the scale of Hilbert spaces $H^s(\Omega) := \mathcal{H}_s$. Hence for $s > t > 0$ we have the Gelfand triplets

$$H^s(\Omega) \hookrightarrow H^s(\Omega) \hookrightarrow H^0(\Omega) \cong L_2(\Omega) \hookrightarrow (H^t(\Omega))^* \hookrightarrow (H^s(\Omega))^*.$$

We then define the closed subspaces $\mathring{H}^s(\Omega) := \mathrm{cl}_{H^s} C_{00}^\infty(\Omega)$, their duals again being denoted by $H^{-s}(\Omega)$. In a similar vein the spaces $W_p^s(\Omega)$ and their closed subspaces $\mathring{W}_p^s(\Omega)$ with norms $\|u\|_{s,p}$ may be defined. The dual of the latter subspace is again denoted by $W_{p^*}^{-s}(\Omega)$.

A space which is occasionally useful is $W_{pu}^k(\Omega) := W_p^k(\Omega) \cap C_u^k(\Omega)$, as there is a dense embedding

$$W_{pu}^k(\Omega) \hookrightarrow W_p^k(\Omega).$$

Let $0 \le t \le s \in \mathbb{R}$, $1 \le p, q < \infty$, and $1 \le m \le n \in \mathbb{N}$, as well as $s - t \ge \frac{n}{p} - \frac{m}{q}$ with $\ell \in \mathbb{N}_0$ the integer part of t, then the following embeddings are continuous and dense:

$$0) \qquad W_p^s(\mathbb{R}^n) \hookrightarrow W_q^\ell(\mathbb{R}^m).$$

In addition, for any subdomain $\Omega \subseteq \mathbb{R}^n$ with locally Lipschitz boundary,

$$\begin{array}{ll} \text{i)} & W_p^s(\Omega) \hookrightarrow W_p^t(\Omega) \hookrightarrow L_p(\Omega), \\ \text{i)} & \mathring{W}_p^s(\Omega) \hookrightarrow \mathring{W}_p^t(\Omega) \hookrightarrow L_p(\Omega), \\ \text{iii)} & \mathscr{S}(\Omega) \hookrightarrow W_p^s(\Omega) \hookrightarrow W_p^t(\Omega), \\ \text{iv)} & \mathscr{D}(\Omega) \hookrightarrow C_{00}^k(\Omega) \hookrightarrow \mathring{W}_p^s(\Omega) \hookrightarrow \mathring{W}_p^t(\Omega). \end{array}$$

For an additionally bounded domain Ω, we also need the spaces $C^k(\bar{\Omega}) = C_b^k(\bar{\Omega}) = C_u^k(\bar{\Omega})$, and $C^{k,\beta}(\bar{\Omega}) = C_b^{k,\beta}(\bar{\Omega}) = C_u^{k,\beta}(\bar{\Omega})$, which are Banach spaces when equipped with the corresponding norms (see section 6.2). For $k < m$ and $0 < \alpha < \beta < 1$, we have the continuous, dense, and compact embeddings:

$$C^{m,1}(\bar{\Omega}) \hookrightarrow C^{m,\beta}(\bar{\Omega}) \hookrightarrow C^{m,\alpha}(\bar{\Omega}) \hookrightarrow C^{m}(\bar{\Omega}) = C^{m}_{u}(\Omega) \hookrightarrow C^{k,1}(\bar{\Omega}).$$

If $1 \le p \le q \le \infty$, $0 \le t \le s \le k \in \mathbb{N}$, and $s - t \ge n\left(\frac{1}{p} - \frac{1}{q}\right)$, the following embeddings are dense and continuous:

$$\begin{array}{ll} \text{v)} & C^{\infty}(\bar{\Omega}) \hookrightarrow C^{k}(\bar{\Omega}) \hookrightarrow W^{s}_{p}(\Omega) \hookrightarrow W^{t}_{p}(\Omega), \\ \text{vi)} & \mathscr{D}(\Omega) \hookrightarrow C^{k}_{00}(\Omega) \hookrightarrow \mathring{W}^{s}_{p}(\Omega) \hookrightarrow \mathring{W}^{t}_{p}(\Omega), \\ \text{vii)} & W^{s}_{p}(\Omega) \hookrightarrow W^{t}_{q}(\Omega), \quad \mathring{W}^{s}_{p}(\Omega) \hookrightarrow \mathring{W}^{t}_{q}(\Omega). \end{array}$$

Furthermore with $0 < \beta < 1$, the following embeddings are continuous and dense:

$$\begin{array}{rlcl} \text{viii) for } \dfrac{n}{p} < s < \dfrac{n+1}{p}: & W^{s}_{p}(\Omega) & \hookrightarrow & C^{0,s-(n/p)}(\bar{\Omega}), \\ \text{ix) for } s = \dfrac{n+1}{p}: & W^{s}_{p}(\Omega) & \hookrightarrow & C^{0,\beta}(\bar{\Omega}), \\ \text{x) for } s > \dfrac{n+1}{p}: & W^{s}_{p}(\Omega) & \hookrightarrow & C^{0,1}(\bar{\Omega}). \end{array}$$

In addition, if $1 \le p, q < \infty$, $0 \le t \le s$, $s - t > n\left(\frac{1}{p} - \frac{1}{q}\right)$, and $\ell \in \mathbb{N}_0$ with $\ell \le s$, the following embeddings are as well compact:

$$\begin{array}{rlcl} \text{xi)} & W^{s}_{p}(\Omega) & \hookrightarrow & W^{t}_{q}(\Omega), \\ \text{xii) for } s - \ell > \dfrac{n}{p}: & W^{s}_{p}(\Omega) & \hookrightarrow & C^{\ell}(\bar{\Omega}) \subset L_{\infty}(\Omega). \end{array}$$

The trace on a smooth r-dimensional submanifold Γ_r inside Ω will be considered, as well as the trace on a smooth part of the boundary $\Gamma_b \subseteq \partial\Omega$.

First consider the internal manifold Γ_r. Let $1 \le p, q < \infty$, $0 \le t < s$, $s - t > \frac{n}{p} - \frac{r}{q}$, and let $k \in \mathbb{N}_0$, $k \le t$, then there is a $C > 0$ such that for $u \in W^{s}_{p}(\Omega)$:

$$\text{xiii)} \qquad \|\mathrm{D}^{k} u\|_{W^{t-k}_{q}(\Gamma_r)} \le C \|u\|_{W^{s}_{p}(\Omega)}.$$

As Γ_b is $(n-1)$-dimensional, the only difference is that one now requires $s - t > \frac{n}{p} - \frac{n-1}{q}$, then for $u \in W^{s}_{p}(\Omega)$:

$$\text{xiv)} \qquad \|\mathrm{D}^{k} u\|_{W^{t-k}_{q}(\Gamma_b)} \le C \|u\|_{W^{s}_{p}(\Omega)}.$$

Both maps onto either Γ_r or Γ_b are compact.

7 Direct Sums

For a collection $\mathcal{H}_n$ of Hilbert or Banach spaces, the product $\prod_n \mathcal{H}_n$ is also a topological space (with the product topology), but is *not normable* in case the collection is infinite, (only metrisable in case of a countable collection analogous to $\boldsymbol{\omega}$ in section 6.1); for the direct sum this is different.

If $\mathcal{H}_n, n \in \mathcal{C}$ is a—possibly infinite but countable—collection of Hilbert spaces, an inner product may defined on $\bigoplus_n \mathcal{H}_n$ by

$$\forall x = (x_n), y = (y_n) \in \bigoplus_{n\in\mathcal{C}} \mathcal{H}_n : \quad \langle x|y\rangle_\oplus := \sum_{n\in\mathcal{C}} \langle x_n|y_n\rangle_n,$$

and $\bigoplus_{(2)\, n\in\mathcal{C}} \mathcal{H}_n$ is the completion of $\bigoplus_{n\in\mathcal{C}} \mathcal{H}_n$ in the induced topology. This means

$$x \in \bigoplus_{n\in\mathcal{C}}{}_{(2)} \mathcal{H}_n \iff x \in \prod_{n\in\mathcal{C}} \mathcal{H}_n \ \wedge \ \left(\sqrt{\langle x_n|x_n\rangle_n}\right) \in \ell_2.$$

If the $\mathcal{H}_n$ are all equal to some Hilbert space $\mathcal{H}$, this is often denoted by $\ell_2(\mathcal{H})$. With the natural injection $\iota_n : \mathcal{H}_n \to \bigoplus_{(2)} \mathcal{H}_n$, one then has for $n \neq m$ that $\iota_n(\mathcal{H}_n) \perp \iota_m(\mathcal{H}_m)$.

Similarly for a collection of Banach spaces $\mathcal{X}_n, n \in \mathcal{C}$, a variety of norms may defined on $\bigoplus_{n\in\mathcal{C}} \mathcal{X}_n$, e.g. for all $1 \leq p < \infty$ by

$$\forall x = (x_n) \in \bigoplus_{n\in\mathcal{C}} \mathcal{X}_n : \quad \|x\|_p^p := \sum_{n\in\mathcal{C}} \|x_n\|_n^p,$$

and $\bigoplus_{(p)\, n\in\mathcal{C}} \mathcal{X}_n$ is the completion of $\bigoplus_{n\in\mathcal{C}} \mathcal{X}_n$ in the induced topology. Hence

$$x \in \bigoplus_{n\in\mathcal{C}}{}_{(p)} \mathcal{X}_n \iff x \in \prod_{n\in\mathcal{C}} \mathcal{X}_n \ \wedge \ (\|x_n\|_n) \in \ell_p.$$

The dual of this space (for $1 < p < \infty$) is $\bigoplus_{(p^*)\, n\in\mathcal{C}} \mathcal{X}_n^*$, with p^* as before. The natural duality pairing between $x = (x_n) \in \bigoplus_{(p)\, n\in\mathcal{C}} \mathcal{X}_n$ and $x^* = (x_n^*) \in \bigoplus_{(p^*)\, n\in\mathcal{C}} \mathcal{X}_n^*$ is $\langle x^*, x\rangle_{\mathcal{X}^*\times\mathcal{X}} := \sum_{n\in\mathcal{C}} \langle x_n^*, x_n\rangle_{\mathcal{X}_n^*\times\mathcal{X}_n}$.

For $p = \infty$ set

$$x \in \bigoplus_{n\in\mathcal{C}}{}_{(\infty)} \mathcal{X}_n \iff x \in \prod_{n\in\mathcal{C}} \mathcal{X}_n \ \wedge \ \|x\|_\infty := \sup_n \|x_n\|_n < \infty,$$

and for $p = 0$ set

$$x \in \bigoplus_{n\in\mathcal{C}}{}_{(0)} \mathcal{X}_n \iff x \in \prod_{n\in\mathcal{C}} \mathcal{X}_n \ \wedge \ (\|x_n\|_n) \in \boldsymbol{c}_0.$$

Note that $(\bigoplus_{(0)\, n\in\mathcal{C}} \mathcal{X}_n)^* \cong \bigoplus_{(1)\, n\in\mathcal{C}} \mathcal{X}_n^*$, and $(\bigoplus_{(1)\, n\in\mathcal{C}} \mathcal{X}_n)^* \cong \bigoplus_{(\infty)\, n\in\mathcal{C}} \mathcal{X}_n^*$.

If the $\mathcal{X}_n$ are all equal to some Banach space $\mathcal{X}$, this is often denoted by $\ell_p(\mathcal{X})$, or $\boldsymbol{c}_0(\mathcal{X})$ respectively.

8 Tensor Products

For two vector spaces $\mathcal{X}$ and $\mathcal{Y}$, the tensor product is denoted by $\mathcal{X}\otimes\mathcal{Y}$. For $(x,y)\in\mathcal{X}\times\mathcal{Y}$, define $x\otimes y\in\mathscr{L}(\mathcal{X}\times\mathcal{Y},\mathbb{K})^* = \mathscr{L}^2(\mathcal{X},\mathcal{Y};\mathbb{K})^*$ by

$$\forall u\in\mathscr{L}(\mathcal{X}\times\mathcal{Y},\mathbb{K}): \quad x\otimes y:\, u\mapsto (x\otimes y)(u) := u(x,y).$$

The map $\otimes:\mathcal{X}\times\mathcal{Y}\to\mathscr{L}^2(\mathcal{X},\mathcal{Y};\mathbb{K})^*$ is bilinear, and set

$$\mathcal{X}\otimes\mathcal{Y} := \operatorname{span}\otimes(\mathcal{X}\times\mathcal{Y})\subseteq\mathscr{L}^2(\mathcal{X},\mathcal{Y};\mathbb{K})^*.$$

This is easily extended to any finite number of spaces $\bigotimes_{k=1}^n \mathcal{X}_k$. If the $\mathcal{X}_k$ are all equal, this is the n-th tensorial power of $\mathcal{X}$, denoted by $\mathcal{X}^{\otimes n} := \bigotimes_{k=1}^n \mathcal{X}$, with elements as linear combination of terms of the form $\bigotimes_{k=1}^n x_k := x_1\otimes\cdots\otimes x_n$. The space and elements are called **homogeneous** tensors of **degree** or **grade** $\deg(x_1\otimes\cdots\otimes x_n) = n$ Observe that $\mathcal{X}^*\otimes\mathcal{Y}\cong\mathscr{L}_{00}(\mathcal{X},\mathcal{Y})$, as $\forall x\in\mathcal{X}:\; x^*\otimes y: x\mapsto\langle x^*,x\rangle\, y\in\mathcal{Y}$ is linear in x.

8.1 Tensor Algebra

For a vector space $\mathcal{X}$, the n-fold **contravariant** and m-fold **covariant** tensor product—of type or degree $\binom{m}{n}$ —is

$$\mathscr{T}_n^m(\mathcal{X}) := \mathcal{X}^{\otimes n}\otimes\mathcal{X}^{*\otimes m}\subseteq\mathscr{L}^{n+m}(\mathcal{X}^{*n},\mathcal{X}^m;\mathbb{R}).$$

The space $\mathscr{T}_n^m(\mathcal{X})\otimes\mathcal{Y}$ is the space of $\mathcal{Y}$-valued tensors of type $\binom{m}{n}$, a subspace of $\mathscr{L}^{n+m}(\mathcal{X}^{*n},\mathcal{X}^m;\mathcal{Y})$.

With $\mathcal{X}^{\otimes 0} := \mathbb{K}$ and $\mathcal{X}^{\otimes 1} := \mathcal{X}$, the tensor algebra is denoted by

$$\mathscr{T}(\mathcal{X}) := \mathbb{K}\oplus\bigoplus_{m,n\in\mathbb{N}}\mathscr{T}_n^m(\mathcal{X}).$$

It is an associative **graded** algebra with the "multiplication" $\mathscr{T}(\mathcal{X})\ni x,y\mapsto x\otimes y$. Observe that $\mathcal{X}\otimes\mathbb{K}\cong\mathcal{X}$.

An important operation on $\mathscr{T}(\mathcal{X})$ is the contraction. If $\jmath\le n, \imath\le m$, the contraction $C_\jmath^\imath:\;\mathscr{T}_n^m(\mathcal{X})\to\mathscr{T}_{n-1}^{m-1}(\mathcal{X})$ is defined by linear extension of

$$C_\jmath^\imath:\; x_1\otimes\cdots x_\jmath\cdots\otimes x_n\otimes x_1^*\otimes\cdots x_\imath^*\cdots\otimes x_m^*\mapsto \langle x_\imath^*, x_\jmath\rangle\, x_1\otimes\cdots\widehat{x_\jmath}\cdots\otimes x_n\otimes x_1^*\otimes\cdots\widehat{x_\imath^*}\cdots\otimes x_m^*,$$

where the "hat" means that this argument is omitted. The contractions C_k^k may be used to show the duality of spaces $\mathscr{T}_n^0(\mathcal{X}) \subseteq \mathscr{L}^n(\mathcal{X};\mathbb{K})^*$ and $\mathscr{T}_0^n(\mathcal{X}) \subseteq \mathscr{L}^n(\mathcal{X};\mathbb{K})$ via linear extension of

$$\left\langle \bigotimes_{k=1}^n x_k^*, \bigotimes_{k=1}^n x_k \right\rangle := C_1^1 \circ \cdots \circ C_n^n \left(\bigotimes_{k=1}^n x_k^* \otimes \bigotimes_{k=1}^n x_k \right) = \prod_{k=1}^n \langle x_k^*, x_k \rangle.$$

Two sub-algebras are $\mathscr{T}_*(\mathcal{X}) := \bigoplus_{n\in\mathbb{N}_0} \mathcal{X}^{\otimes n}$, also denoted as $\bigotimes(\mathcal{X})$, the contravariant tensors, and the covariant tensors $\mathscr{T}^*(\mathcal{X}) := \bigoplus_{m\in\mathbb{N}_0} \mathcal{X}^{*\otimes m}$, also denoted as $\bigotimes(\mathcal{X}^*)$. If $B \in \mathscr{L}(\mathcal{X})$, this may be extended to each $\mathcal{X}^{\otimes n}$ via

$$\Gamma_n(B) := B^{\otimes n} : \bigotimes_{k=1}^n x_k \mapsto \bigotimes_{k=1}^n B(x_k),$$

and thereafter to all of $\mathscr{T}_*(\mathcal{X})$ via $\Gamma(B) := B^{\otimes} := \bigoplus_{n\in\mathbb{N}_0} B^{\otimes n}$. On $\mathcal{X}^{\otimes n}$ note

$$\mathrm{d}\Gamma_n(B) : \bigotimes_{k=1}^n x_k \mapsto \sum_{k=1}^n x_1 \otimes \cdots \otimes B(x_k) \otimes \cdots \otimes x_n,$$

and its extension to all of $\bigotimes(\mathcal{X})$ via $\mathrm{d}\Gamma(B) := \bigoplus_{n\in\mathbb{N}_0} \mathrm{d}\Gamma_n(B)$.

The spaces $\mathscr{T}_*(\mathcal{X}) \otimes \mathcal{Y}$, $\mathscr{T}^*(\mathcal{X}) \otimes \mathcal{Y}$, and $\mathscr{T}(\mathcal{X}) \otimes \mathcal{Y}$ are the contravariant, covariant, or mixed $\mathcal{Y}$ valued tensors.

8.2 Symmetric Tensors

Let $y \in \mathscr{T}_n$ and $x \in \mathscr{T}_m$ be such that $S_n(y) = y$ and $S_m(z) = z$ (section 4.4), i.e. both are *symmetric tensors*. Then the **symmetric tensor product** of y and z is given by

$$y \vee z := \frac{(n+m)!}{n!\,m!} S_{(n+m)}(y \otimes z).$$

A completely analogous definition holds for elements in $\mathscr{T}^n$ and $\mathscr{T}^m$.

The subspace of $\mathcal{X}^{\otimes n}$ generated by the symmetric tensors is denoted by $\mathcal{X}^{\vee n} := \bigvee_{k=1}^n \mathcal{X}$. The symmetric tensor algebra $\bigvee(\mathcal{X}) := \bigoplus_{k\in\mathbb{N}_0} \mathcal{X}^{\vee k}$ is the **symmetric Fock space**. It is an associative Abelian graded algebra with the "multiplication" $\bigvee(\mathcal{X}) \ni x, y \mapsto x \vee y = y \vee x$. The space $\bigvee(\mathcal{X}) \otimes \mathcal{Y}$ is the space of symmetric $\mathcal{Y}$-valued tensors. Alternatively, consider in $\mathscr{T}_*(\mathcal{X})$ the ideal $\mathcal{I}$ generated by the expressions of the form $x \otimes y - y \otimes x$. Then $\bigvee(\mathcal{X})$ is equally well described as the factor algebra $\mathscr{T}_*(\mathcal{X})/\mathcal{I}$.

If $B \in \mathscr{L}(\mathcal{X})$, this may be extended to each $\mathcal{X}^{\vee n}$ via

$$B^{\vee n} : \bigvee_{k=1}^{n} x_k \mapsto \bigvee_{k=1}^{n} B(x_k),$$

and thereafter to all of $\bigvee(\mathcal{X})$ via $\Gamma(B) := \bigoplus_{n\in\mathbb{N}_0} B^{\vee n}$. Note also on $\mathcal{X}^{\vee n}$

$$\mathrm{d}\Gamma_n(B) : \bigvee_{k=1}^{n} x_k \mapsto \sum_{k=1}^{n} x_1 \vee \cdots \vee B(x_k) \vee \cdots \vee x_n,$$

and its extension to all of $\bigvee(\mathcal{X})$ via $\mathrm{d}\Gamma(B) := \bigoplus_{n\in\mathbb{N}_0} \mathrm{d}\Gamma_n(B)$.

8.3 Alternating or Anti-Symmetric Tensors or Forms

This development is in analogy to the symmetric case in section 8.2. Let $y \in \mathscr{T}_n$ and $x \in \mathscr{T}_m$ be such that $A_n(y) = y$ and $A_m(z) = z$ (section 4.4), i.e. both are *skew-symmetric tensors*. Then the **alternating** or **anti-symmetric tensor product** of y and z is given by

$$y \wedge z := \frac{(n+m)!}{n!\, m!} A_{(n+m)}(y \otimes z).$$

A completely analogous definition holds for elements in $\mathscr{T}^n$ and $\mathscr{T}^m$. This product is also called the **exterior** or **Grassmann product**. Note that for $x \in \mathcal{X} : x \wedge x = 0$. The subspace of $\mathcal{X}^{\otimes n}$ generated by these anti-symmetric tensor products is denoted by $\mathcal{X}^{\wedge n}$; elements of the form $\bigwedge_{k=1}^{n} x_k$ are called **n-vectors** or sometimes **blades**.

The anti-symmetric tensor product is often only considered as a subspace of $\mathcal{X}^{*\otimes n}$. The subspace of these anti-symmetric tensor products is denoted by $\Lambda^n(\mathcal{X}) := \mathcal{X}^{*\wedge n} := \bigwedge_{k=1}^{n} \mathcal{X}^*$.

The anti-symmetric tensor algebra, or exterior algebra, or Grassmann algebra is $\Lambda(\mathcal{X}) := \bigoplus_{k\in\mathbb{N}_0} \Lambda^k(\mathcal{X})$, also called the **anti-symmetric Fock space**. It is also called the space of **forms**. It is an associative graded algebra with the "multiplication" $\Lambda(\mathcal{X}) \ni x^*, y^* \mapsto x^* \wedge y^*$. If $x^* \in \Lambda^k(\mathcal{X})$ is homogeneous with $\deg x^* = k$ and $y^* \in \Lambda^m(\mathcal{X})$ is homogeneous with $\deg y^* = m$, then $x^* \wedge y^* = (-1)^{km} y^* \wedge x^*$. The space $\Lambda(\mathcal{X}) \otimes \mathcal{Y}$ is the space of $\mathcal{Y}$-valued forms. Alternatively, consider in $\mathscr{T}^*(\mathcal{X})$ the ideal $\mathcal{I}$ generated by the expressions of the form $x^* \otimes x^*$. Then $\Lambda(\mathcal{X})$ is equally well described as the factor algebra $\mathscr{T}^*(\mathcal{X})/\mathcal{I}$.

One important operation on $\Lambda^k(\mathcal{X})$ is the **inner product** with a vector $v \in \mathcal{X}$, a mapping $\boldsymbol{\iota}_v : \Lambda^k(\mathcal{X}) \to \Lambda^{k-1}(\mathcal{X})$ defined by

$$\boldsymbol{\iota}_v : x_1^* \wedge \cdots \wedge x_k^* \mapsto \sum_{m=1}^{k} (-1)^{(m+1)} \langle x_m^*, v\rangle \, x_1^* \wedge \cdots \widehat{x_m^*} \cdots \wedge x_k^*.$$

This map is an **anti-derivation**, i.e. if $x^* \in \Lambda^k(\mathcal{X})$ is homogeneous with $\deg x^* = k$, then $\boldsymbol{\iota}_v(x^* \wedge y^*) = (\boldsymbol{\iota}_v x^*) \wedge y^* + (-1)^k x^* \wedge (\boldsymbol{\iota}_v y^*)$.

8.4 Tensor Product of Hilbert Spaces

If $\mathcal{H}$ and $\mathcal{K}$ are Hilbert spaces, a new inner product may be defined on $\mathcal{H} \otimes \mathcal{K}$ by

$$\forall u \otimes v, x \otimes y \in \mathcal{H} \otimes \mathcal{K} : \quad \langle u \otimes v | x \otimes y \rangle_{\otimes} := \langle u | x \rangle_{\mathcal{H}} \langle v | y \rangle_{\mathcal{K}}.$$

Observe that for the induced norm on $\mathcal{H} \otimes \mathcal{K}$ the **cross-norm** property holds: $\|x \otimes y\|_{\mathcal{H} \otimes \mathcal{K}} = \|x\|_{\mathcal{H}} \|y\|_{\mathcal{K}}$.

The completion of $\mathcal{H} \otimes \mathcal{K}$ in the induced metric is again denoted $\mathcal{H} \otimes \mathcal{K}$, but sometimes also as $\mathcal{H} \otimes_{(2)} \mathcal{K}$ or $\mathcal{H} \hat{\otimes} \mathcal{K}$, or even $\mathcal{H} \hat{\otimes}_{(2)} \mathcal{K}$. This inner product can be extended to the tensor algebra $\mathscr{T}(\mathcal{H})$, as well as to the subspaces $\Gamma(\mathcal{H})$ and $\Lambda(\mathcal{H})$, which are orthogonal to each other. Observe that $\mathcal{H}^* \otimes \mathcal{H} \cong \mathscr{L}_2(\mathcal{H})$, the space of *Hilbert-Schmidt operators*, and that $\mathcal{H}^* \vee \mathcal{H}$ is the space of *symmetric* Hilbert-Schmidt operators.

Note also that $L_2(\Omega; \mathcal{H}) \cong L_2(\Omega) \otimes \mathcal{H}$. Especially $L_2(\Omega_1 \times \Omega_2) \cong L_2(\Omega_1; L_2(\Omega_2)) \cong L_2(\Omega_2; L_2(\Omega_1)) \cong L_2(\Omega_1) \otimes L_2(\Omega_2)$.

8.5 Tensor Products of Banach Spaces

There is no "canonical" way as in section 8.4 to assign a norm on the tensor product $\mathcal{X} \otimes \mathcal{Y}$ of two Banach spaces $\mathcal{X}$ and $\mathcal{Y}$.

One way is to observe that $\mathcal{X} \otimes \mathcal{Y} \subseteq \mathscr{L}(\mathcal{X}^*, \mathcal{Y})$, and the latter is a natural Banach space. This induces a norm on the subspace, and the completion in that norm is denoted by $\mathcal{X} \otimes_\varepsilon \mathcal{Y}$ or $\mathcal{X} \otimes_{(\infty)} \mathcal{Y}$ (or even $\mathcal{X} \hat{\otimes}_\varepsilon \mathcal{Y}$), and the norm—which is just the operator norm—by $\| \cdot \|_\varepsilon$ or sometimes by $\| \cdot \|_{(\infty)}$, as it is also a completion in the Schatten-∞ norm (see section 4.6.1), i.e.

$$\|x \otimes y\|_\varepsilon := \|x \otimes y\|_{(\infty)} := \sup_{\|x^*\|_{\mathcal{X}^*} = 1} \|\langle x^*, x \rangle y\|_{\mathcal{Y}} = \|x\|_{\mathcal{X}} \|y\|_{\mathcal{Y}}.$$

This means obviously that this norm satisfies the *cross-norm property*. The tensor product with this topology is called the **injective tensor product**, and the norm sometimes the λ-norm, as it is the least norm—giving the weakest topology—which satisfies the cross-norm property, and is then denoted as $\| \cdot \|_\lambda$, and the tensor product with $\otimes_\lambda$. If the so-called *approximation property* is satisfied (compact operators may be approximated in norm by finite-rank operators—but this is not true for all Banach spaces), then obviously the injective tensor product will coincide with the space of *compact operators* from $\mathcal{X}^*$ to $\mathcal{Y}$.

Observe that for any topological spaces S and T, one has $C(\mathsf{S}\times\mathsf{T}) \cong C(\mathsf{S}, C(\mathsf{T})) \cong C(\mathsf{T}, C(\mathsf{S})) \cong C(\mathsf{S}) \otimes_\varepsilon C(\mathsf{T})$. In fact, for any Banach space $\mathcal{X}$ one has $C(\mathsf{S}, \mathcal{X}) \cong C(\mathsf{S}) \otimes_\varepsilon \mathcal{X}$.

Another complementary way is for $z \in \mathcal{X} \otimes \mathcal{Y}$ to set

$$\|z\|_\pi = \|z\|_{(1)} := \inf \left\{ \sum_j \|x_j\|_{\mathcal{X}}\ \|y_j\|_{\mathcal{Y}} : \quad z = \sum_j x_j \otimes y_j \right\},$$

where the infimum is over all possible representations of $z \in \mathcal{X} \otimes \mathcal{Y}$. The completion of $\mathcal{X} \otimes \mathcal{Y}$ with the metric induced by this norm is denoted by $\mathcal{X} \otimes_\pi \mathcal{Y}$ or $\mathcal{X} \otimes_{(1)} \mathcal{Y}$ (or even $\mathcal{X} \hat{\otimes}_\pi \mathcal{Y}$).

This norm also satisfies the *cross-norm property*. The tensor product with this topology is called the **projective tensor product**, and the norm sometimes the γ-norm, as it is the greatest norm—giving the strongest topology—which satisfies the cross-norm property, and is then denoted as $\|\cdot\|_\gamma$, and the tensor product with $\otimes_\gamma$. Note that $\|x \otimes y\|_\varepsilon \leq \|x \otimes y\|_\pi$. In case the approximation property holds, the projective tensor product is just the completion in the Schatten-1 norm, the space of *nuclear maps*. Note that completions in all other Schatten-p norms would be possible, giving other spaces with the cross-norm property (see section 4.6.1). These tensor products can then be denoted as $\otimes_{(p)}$ and the tensor product norms with $\|\cdot\|_{(p)}$.

Observe that for measure spaces S and T, one has that $L_1(\mathsf{S}\times\mathsf{T}) \cong L_1(\mathsf{S}, L_1(\mathsf{T})) \cong L_1(\mathsf{T}, L_1(\mathsf{S})) \cong L_1(\mathsf{S}) \otimes_\pi L_1(\mathsf{T})$. In fact, for any Banach space $\mathcal{X}$ one has $L_1(\mathsf{S}; \mathcal{X}) \cong L_1(\mathsf{S}) \otimes_\pi \mathcal{X}$. Note also that $(\mathcal{X} \otimes_\pi \mathcal{Y})^* \cong \mathscr{L}(\mathcal{X}, \mathcal{Y}^*)$.

9 Different Vector Space Notations

If there are conceptually different vector spaces involved, it is prudent to keep them notationally apart by e.g. using for one type $\mathcal{X}$ for the spaces, A for mappings, and $x \in \mathcal{X}$ for elements. Corresponding capital *Greek letters* like $\mathit{\Gamma}, \mathit{\Lambda}, \mathit{\Xi}, \mathit{\Omega}$.

The next type may be denoted by using $\boldsymbol{\mathcal{X}}$ for the spaces, $\boldsymbol{A}$ for mappings, and $\boldsymbol{x} \in \boldsymbol{\mathcal{X}}$ for elements, and capital *Greek letters* like $\boldsymbol{\mathit{\Gamma}}, \boldsymbol{\mathit{\Lambda}}, \boldsymbol{\mathit{\Xi}}, \boldsymbol{\mathit{\Omega}}$.

The next type might be denoted by $\mathfrak{X}$ for the spaces, A for mappings, and $\mathsf{x} \in \mathfrak{X}$ for elements, and capital *Greek letters* like $\Gamma, \Lambda, \Xi, \Omega$.

Again the next type may be denoted by using $\boldsymbol{\mathfrak{X}}$ for the spaces, $\boldsymbol{\mathsf{A}}$ for mappings, and $\boldsymbol{\mathsf{x}} \in \boldsymbol{\mathfrak{X}}$ for elements, and capital *Greek letters* like $\mathbf{\Gamma}, \mathbf{\Lambda}, \mathbf{\Xi}, \mathbf{\Omega}$.

And again the next notation may be $\mathbb{X}$ for the spaces, $\mathbf{A}$ for mappings, and $\mathbf{x} \in \mathbb{X}$ for elements.

As an example, assume that the elements of an abstract vector space $\mathcal{X}$ have been denoted by x, and linear mappings by A. After choosing bases, the "normal" coordinate vectors in $\mathbb{K}^n$ may be denoted by e.g. $\boldsymbol{x}$, and the corresponding matrices by $\boldsymbol{A}$. If now these are used to build **block** vectors and matrices, these could be denoted by e.g. $\mathbf{x}$ and $\mathbf{A}$.

10 Analysis

Mappings—linear or non-linear—are preferably denoted by *upper* case letters, whereas functionals—i.e. mappings into $\mathbb{K}$ —are often denoted by *lower* case letters. The spaces are denoted in *calligraphic* font, as are subsets. Elements of the vector space are written by *lower* case italic letters (normal math font), and scalars are denoted by *lower* case Greek letters, unless convention dictates some other notation.

10.1 Gâteaux Derivative and Gradient

Let $F \in \mathscr{F}(\mathcal{X}, \mathcal{Y})$ be a mapping, and $\mathcal{X}$ and $\mathcal{Y}$ Banach spaces. The Gâteaux differential at $x \in \mathcal{X}$ in the direction $v \in \mathcal{X}$ is denoted by $\delta F(x, v) := \lim_{\vartheta \to 0}(F(x + \vartheta v) - F(x))/\vartheta$, it is a directional derivative. Note that no topology on $\mathcal{X}$ is required for the Gâteaux differential. But if this is linear and continuous in v —i.e. $\exists \delta F(x) \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$ such that $\delta F(x)v = \delta F(x, v)$— it is called the Gâteaux derivative, and if additionally $\mathcal{Y} = \mathbb{K}$ — i.e. F is a functional—one may write $\langle \delta F(x), v \rangle$. Observe that in this case $\delta F(x) \in \mathcal{X}^*$.

In case $\mathcal{X}$ is a Hilbert space, one may use the inner product to define the gradient $\nabla F(x)$ via $\langle \nabla F(x) | v \rangle_{\mathcal{X}} = \langle \delta F(x), v \rangle \ \forall v \in \mathcal{X}$. In this case $\nabla F(x) \in \mathcal{X}$, whereas $\delta F(x) \in \mathcal{X}^*$.

10.2 Fréchet Derivative

Similarly, the Fréchet derivative with the same conditions is denoted by $\mathrm{D}F(x)v$, where $\mathrm{D}F(x) \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$, and in case $\mathcal{Y} = \mathbb{K}$ by $\langle \mathrm{D}F(x), v \rangle$ with $\mathrm{D}F(x) \in \mathscr{L}(\mathcal{X}, \mathbb{K}) = \mathcal{X}^*$. If both the Gâteaux and Fréchet derivatives exist, they coincide.

A map $F \in \mathscr{F}(\mathcal{X}, \mathcal{Y})$ is called positively homogeneous of degree k, iff for $\alpha \in \mathbb{R}_+$ one has that $F(\alpha x) = \alpha^k F(x)$. If F is Fréchet differentiable, then one has Euler's formula $\langle \mathrm{D}F(x), x \rangle = kF(x)$.

The k-th order Fréchet derivative of $F \in C^k(\mathcal{X}, \mathcal{Y})$ may be regarded as a symmetric multi-linear map $\mathrm{D}^k F(x) \in \mathscr{L}^k(\mathcal{X}; \mathcal{Y})$. Hence it may be linearised by $\bigvee$ and regarded as a linear map $\mathrm{D}^k F(x) \in \mathscr{L}(\mathcal{X}^{\vee k}, \mathcal{Y})$. In toto, on the whole space $\mathrm{D}^k F \in C(\mathcal{X}, \mathscr{L}(\mathcal{X}^{\vee k}, \mathcal{Y}))$. If for $F \in C^1(\mathcal{X}, \mathcal{X}^*)$ one has that $\mathrm{D}F(x) \in \mathscr{L}(\mathcal{X}; \mathcal{X}^*)$ is self-adjoint (see section 4.6) for all $x \in \mathcal{X}$, then F is called a **gradient** operator; i.e. there is a functional $\phi \in C^2(\mathcal{X}, \mathbb{K})$, such that $F = \mathrm{D}\phi$. The operator "D" is a derivation—as $\mathrm{D}(\mathsf{F} \otimes \mathsf{G}) = (\mathrm{D}\mathsf{F}) \otimes \mathsf{G} + \mathsf{F} \otimes (\mathrm{D}\mathsf{G})$ —on the algebra of tensor valued maps.

In any case, Taylor's formula reads

$$F(x+h) = \sum_{k \geq 0} \frac{1}{k!} \mathrm{D}^k F(x) h^{\vee k}.$$

10.3 Exterior Derivative

For $F \in C^1(\mathcal{X}, \mathcal{Z})$ with $\mathcal{Z} = \mathscr{L}(\mathcal{X}^{\wedge k}, \mathcal{Y}) \subseteq \mathscr{L}^k(\mathcal{X}; \mathcal{Y})$, i.e. an $\mathcal{Y}$-valued form of grade k, the alternating or anti-symmetric part of $\mathrm{D}F(x) \in \mathscr{L}^{(k+1)}(\mathcal{X}; \mathcal{Y})$ is the exterior derivative $\mathrm{d}F(x) := A_{k+1}(\mathrm{D}F(x)) \in \mathscr{L}(\mathcal{X}^{\wedge(k+1)}, \mathcal{Y})$, a form of grade $k+1$. In analogy to section 10.2, in toto $\mathrm{d}F \in C(\mathcal{X}, \mathscr{L}(\mathcal{X}^{\wedge(k+1)}, \mathcal{Y}))$. Note that for $f \in C^1(\mathcal{X}, \mathbb{K})$ (i.e. $k = 0$ and $\mathcal{Y} = \mathbb{R}$), one has $\mathrm{D}f = \mathrm{d}f$, as $\mathcal{X}^{*\wedge 1} = \mathcal{X}^{*\vee 1} = \mathcal{X}^*$. Observe that $\mathrm{d}^2 f := \mathrm{d} \circ \mathrm{d}\, f := \mathrm{d}(\mathrm{d}\, f) = 0$ for any f. The operator "d" is an anti-derivation in the algebra of $\Lambda(\mathcal{X})$ valued maps, i.e. if the values of F are in $\mathcal{X}^{*\wedge k}$, then $\mathrm{d}(\mathsf{F} \wedge \mathsf{G}) = (\mathrm{d}\mathsf{F}) \wedge \mathsf{G} + (-1)^k \mathsf{F} \wedge (\mathrm{d}\mathsf{G})$.

10.4 Integration of Vector Valued Functions

The Lebesgue integral for real valued functions on a measure space $(\Omega, \mathfrak{A}, \mu)$ —see section 3 —is presumed to be known, and also the Lebesgue spaces of integrable functions—see section 6.3.

For a measure space $(\Omega, \mathfrak{A}, \mu)$, a LCS $\mathcal{X}$, a vector valued function $f \in \mathscr{F}(\Omega, \mathcal{X})$, and any $x^* \in \mathcal{X}^*$, define the function $x^* f : \Omega \ni \omega \mapsto x^* f(\omega) := \langle x^*, f(\omega) \rangle \in \mathbb{R}$. If this new function is measurable for all $x^* \in \mathcal{X}^*$, then f is called **weakly μ-measurable** or just weakly measurable.

Maps $f \in \mathscr{F}(\Omega, \mathcal{X})$, which assume only finitely many values $\{x_j\}_{j=1,\dots,m}$, and where the sets $E_j \in \mathfrak{A}$ in Ω where f is constant— $f_{E_j} \equiv x_j$ —are measurable, are called *simple functions*. Their integral over any measurable set A is defined in the obvious way as $\mathcal{X} \ni v_A = \int_A f \, \mathrm{d}\mu := \sum_{j=1}^m x_j \mu(A \cap E_j)$. Assume that $\mathcal{X}$ is a Banach space. Then pointwise norm-limits of such simple functions are called **strongly μ-measurable**, and their **Bochner integral** is the limit of the integrals over the simple functions. Note that in this case the function $\|f\|_{\mathcal{X}}$ is also measurable.

Finally, if $f^{-1}(B) \in \mathfrak{A}$ for any $B \in \mathfrak{B}_{\mathcal{X}}$ (the Borel algebra of $\mathcal{X}$), f is called **Borel μ-measurable**. The map f is called **μ-essentially separably valued**, if $f(\Omega \setminus N)$ is separable, where $N \in \mathfrak{A}$ is a null-set. Observe that a mapping which is strongly μ-measurable, is equivalently weakly or Borel μ-measurable and μ-essentially separably valued. One may write $L_0(\Omega, \mathfrak{A}, \mu; \mathcal{X})$ or just $L_0(\Omega; \mathcal{X})$ for such (equivalence classes of) measurable maps, and similarly $L_p(\Omega; \mathcal{X})$ for $p > 0$ if $\omega \mapsto \|f(\omega)\|_{\mathcal{X}}$ is in the "scalar" $L_p(\Omega)$. Such a function f is therefore **Bochner integrable** iff the function $\Omega \ni \omega \mapsto \|f(\omega\|_{\mathcal{X}} \in \mathbb{R}$ is *Lebesgue integrable*. The vector $\mathcal{X} \ni v_A := \int_A f \,\mathrm{d}\mu$ is called the **Bochner integral** of f over A. Note that $L_1(\Omega; \mathcal{X}) \cong L_1(\Omega) \otimes_\pi \mathcal{X} = L_1(\Omega) \otimes_{(1)} \mathcal{X}$, see section 8.5. Analogously one may define $L_p(\Omega; \mathcal{X}) \cong L_p(\Omega) \otimes_{(p)} \mathcal{X}$ with $1 \leq p \leq \infty$ in the obvious way. Observe that for $1 \leq p < \infty$ one has $L_p(\Omega; \mathcal{X})^* \cong L_{p^*}(\Omega; \mathcal{X}^*)$ where p^* has the usual meaning, if $\mathcal{X}$ has the *Radon-Nikodým property*, for example if $\mathcal{X}$ is reflexive.

A *weakly measurable* function $f \in (\Omega \to \mathcal{X})$ is called **Dunford integrable**, if for all $x^* \in \mathcal{X}^*$ the real valued function $x^* f$ is *Lebesgue integrable* ($x^* f \in L_1(\Omega, \mathfrak{A}, \mu)$). Then there exists a vector $v_A \in \mathcal{X}^{**}$ such that $\forall x^* \in \mathcal{X}^*$ it holds that $\langle x^*, v_A \rangle = \int_A x^* f \,\mathrm{d}\mu$, the vector v_A is called the **Dunford integral** of f over A. In case f is also Bochner integrable, the two integrals coincide. In case the *Dunford integral* satisfies $v_A \in \mathcal{X} \subseteq \mathcal{X}^{**}$, the function is called **Pettis integrable** and v_A is the *Pettis integral*. If $\mathcal{X}$ is a *normed space*, a norm may be defined via $\|f\|_{P_1} = \sup_{x^* \in \mathcal{X}^*} \{ \int x^* f \,\mathrm{d}\mu : \ \|x^*\| \leq 1 \}$. The space of such (equivalence classes of) *Pettis integrable* functions is sometimes denoted by $P_1(\Omega; \mathcal{X})$. Observe that $P_1(\Omega; \mathcal{X}) \cong L_1(\Omega) \otimes_\varepsilon \mathcal{X}$, see section 8.5.

In total we have: Bochner integrable $\Rightarrow$ Pettis integrable $\Rightarrow$ Dunford integrable. Note that the integrals coincide if they all exist. Observe also that Pettis and Dunford integrability do not require that $\mathcal{X}$ be a normed space.

11 Manifolds

If $\mathcal{M}$ is a differentiable manifold, denote by $\boldsymbol{T}_x\mathcal{M}$ the **tangent space** at $x \in \mathcal{M}$, and by $\boldsymbol{T}^*_x\mathcal{M}$ the **cotangent space** at $x \in \mathcal{M}$. If at $x \in \mathcal{M}$ the neighbourhood $\mathcal{U} \in \mathfrak{U}(x)$ provides a local chart, the **tangent bundle** $\boldsymbol{T}(\mathcal{M})$ looks locally like $\biguplus_{x \in \mathcal{U}} \{x\} \times \boldsymbol{T}_x\mathcal{M}$, such that the smooth **bundle section** $\mathbf{v} : \mathcal{U} \ni x \mapsto (x, \boldsymbol{v}(x)) \in \{x\} \times \boldsymbol{T}_x\mathcal{M} \subset \boldsymbol{T}(\mathcal{M})$, is a **vector field**. The local projections for the tangent bundle $\boldsymbol{T}(\mathcal{M})$ and the cotangent bundle $\boldsymbol{T}^*(\mathcal{M})$ (and all other bundles) will be denoted by $\tau_{\mathcal{M}}$, such that $\tau_{\mathcal{M}} : \{x\} \times \boldsymbol{T}_x\mathcal{M} \mapsto x \in \mathcal{M}$. For any section $\mathbf{v}$ in a bundle, it holds that $\tau_{\mathcal{M}} \circ \mathbf{v}(x) = x \in \mathcal{M}$. The space of vector fields on $\mathcal{M}$ is denoted by $\mathscr{X}(\mathcal{M})$, and the space of smooth sections in the cotangent bundle—the covector fields—by $\mathscr{X}^*(\mathcal{M})$; a space of

differential forms of grade 1, i.e. of **Pfaffians**. The smooth bundle sections $\mathscr{E}(\mathcal{M})$ are simply the C^∞ maps into $\mathbb{R}$.

The tangent bundle $\boldsymbol{T}(\mathcal{M}) := \mathscr{T}_1^0(\mathcal{M})$ has as a typical **fibre** $\boldsymbol{T}_x\mathcal{M}$, the cotangent bundle $\boldsymbol{T}^*(\mathcal{M}) := \mathscr{T}_0^1(\mathcal{M})$ has typical fibre $\boldsymbol{T}_x^*\mathcal{M}$. Other tensor bundles may be defined, and the bundle $\mathscr{T}_n^m(\mathcal{M})$ has typical fibre $\mathscr{T}_n^m(\boldsymbol{T}_x\mathcal{M})$ —see section 8.1. Specifically, $\boldsymbol{\Lambda}(\mathcal{M})$ has typical fibre $\Lambda(\boldsymbol{T}_x\mathcal{M})$, i.e. the sections are differential forms. All these bundles are easily given a vector space structure by locally defining for $x \in \mathcal{M}$, $\alpha \in \mathbb{R}$, and $\mathbf{t}, \mathbf{s} \in \mathscr{T}_n^m(\mathcal{M})$ with $\mathbf{t}|_x := (x, \mathsf{t}_x)$ with $\mathsf{t}_x \in \mathscr{T}_n^m(\boldsymbol{T}_x\mathcal{M})$ by $(\alpha\,\mathsf{t})|_x := (x, \alpha\,\mathsf{t}_x)$ for $\alpha\,\mathbf{t}$, and $(\mathsf{s}+\mathsf{t})|_x := (x, \mathsf{s}_x + \mathsf{t}_x)$ for $\mathbf{s}+\mathbf{t}$.

Observe that any space of these vector space valued sections $\{q | q : \mathcal{M} \to \mathscr{T}_n^m(\mathcal{M})$ is not only as usual a vector space by pointwise definition of addition and multiplication of scalars, but also a *module* over $\mathscr{E}(\mathcal{M})$, i.e. by defining the multiplication by a $\phi \in \mathscr{E}(\mathcal{M})$ again pointwise.

Through the duality of tangent and cotangent space, a section $\mathbf{v} \in \mathscr{X}(\mathcal{M})$ in the tangent and a section $\boldsymbol{\beta} \in \mathscr{X}^*(\mathcal{M})$ in the cotangent bundle may be combined to give a real valued function: $\mathscr{E}(\mathcal{M}) \ni \langle \boldsymbol{\beta}, \mathbf{v} \rangle : \mathcal{M} \ni x \mapsto \langle \boldsymbol{\beta}, \mathbf{v} \rangle(x) := \langle \boldsymbol{\beta}(x), \boldsymbol{v}(x) \rangle \in \mathbb{R}$.

11.1 Tangent Map

If $\mathcal{S}$ is another manifold, and $F : \mathcal{M} \to \mathcal{S}$ is a differentiable map, the **tangent map** maps the tangent bundles $\mathbf{T}F \in \mathscr{L}(\boldsymbol{T}(\mathcal{M}), \boldsymbol{T}(\mathcal{S}))$, such that $F \circ \tau_{\mathcal{M}} = \tau_{\mathcal{S}} \circ F_*$, and is locally given by

$$\mathbf{T}F : \boldsymbol{T}(\mathcal{M}) \ni \mathbf{v}(x) = (x, \boldsymbol{v}(x)) \mapsto (F(x), \mathrm{D}F(x)\boldsymbol{v}(x)) \in \boldsymbol{T}(\mathcal{S}).$$

A diffeomorphism is a differentiable map F such that the inverse F^{-1} is also a differentiable map.

As vector bundles are vector spaces under point-wise addition of tangent vectors or differential forms and multiplication by scalars, $\mathbf{T}F$ has a natural dual map, $(\mathbf{T}F)^*$.

In what follows, consider differentiable functions $\varphi, \phi, \in \mathscr{E}(\mathcal{M})$ and $\psi \in \mathscr{E}(\mathcal{S})$, vector fields $\mathbf{u}, \mathbf{v}, \mathbf{y} \in \mathscr{X}(\mathcal{M})$ and $\mathbf{w} \in \mathscr{X}(\mathcal{S})$, as well as two co-vector fields or **differential forms** $\boldsymbol{\beta} \in \mathscr{X}^*(\mathcal{M})$ and $\boldsymbol{\gamma} \in \mathscr{X}^*(\mathcal{S})$.

11.2 Lie Derivative

Both the notion of derivative and exterior derivative of functions can be extended to the manifold and tangent manifold through the device of *directional derivative* in the direction of a *vector* in the *tangent space* $\boldsymbol{T}_x\mathcal{M}$. This involves looking at difference quotients along *integral curves* with this vector as tangent vector. Then for $\varphi \in \mathscr{E}(\mathcal{M})$ one has $\mathrm{d}\varphi \in \mathscr{X}^*(\mathcal{M})$. With a vector field $\mathbf{v} \in \mathscr{X}(\mathcal{M})$ one can associate the directional derivative of φ, the **Lie** derivative, defined by $\mathbf{L_v}\varphi(x) := \langle \mathrm{d}\varphi(x), \mathbf{v}(x)\rangle = \boldsymbol{\iota_v}\mathrm{d}\varphi(x)$ —see section 8.3 for $\boldsymbol{\iota_v}$. It is a **derivation** on the algebra $\mathscr{E}(\mathcal{M})$, as

$$\begin{aligned}\mathbf{L_v}(\varphi(x)\phi(x)) &:= \langle \mathrm{d}(\varphi(x)\phi(x)), \mathbf{v}(x)\rangle \\ &= \phi(x)\langle \mathrm{d}\varphi(x), \mathbf{v}(x)\rangle + \varphi(x)\langle \mathrm{d}\phi(x), \mathbf{v}(x)\rangle \\ &=: \phi(x)\mathbf{L_v}\varphi(x) + \varphi(x)\mathbf{L_v}\phi(x).\end{aligned}$$

Thus the Lie derivative assigns to each differentiable function $\varphi \in \mathscr{E}(\mathcal{M})$ a new function $\mathbf{L_v}\varphi \in \mathscr{E}(\mathcal{M})$, i.e. $\mathbf{L_v} \in \mathscr{L}(\mathscr{E}(\mathcal{M}))$. This allows one to define the commutator $[\mathbf{L_v}, \mathbf{L_y}] := \mathbf{L_v}\circ\mathbf{L_y} - \mathbf{L_y}\circ\mathbf{L_v}$, which is again a derivation. The unique vector field $\mathbf{u}$ such that $[\mathbf{L_v}, \mathbf{L_y}] = \mathbf{L_u}$ will be denoted by $[\mathbf{v}, \mathbf{y}] := \mathbf{u}$, and is called the **Lie bracket** of $\mathbf{v}$ and $\mathbf{y}$, or the *Lie* derivative of $\mathbf{y}$ w.r.t $\mathbf{v}$, also denoted by $\mathbf{L_v y} := [\mathbf{v}, \mathbf{y}]$. With this operation the space of vector fields $\mathscr{X}(\mathcal{M})$ becomes a Lie algebra, as $[\mathbf{v}, \mathbf{y}] = -[\mathbf{y}, \mathbf{v}]$.

For a differential form $\boldsymbol{\beta} \in \mathscr{X}^*(\mathcal{M})$, it follows from the derivation property that its Lie derivative $\mathbf{L_v}\boldsymbol{\beta} \in \mathscr{X}^*(\mathcal{M})$ is the unique differential form such that

$$\forall \mathbf{u} \in \mathscr{X}(\mathcal{M}):\ \langle \mathbf{L_v}\boldsymbol{\beta}, \mathbf{u}\rangle = \mathbf{L_v}(\langle \boldsymbol{\beta}, \mathbf{u}\rangle) - \langle \boldsymbol{\beta}, \mathbf{L_v u}\rangle = \mathbf{L_v}(\langle \boldsymbol{\beta}, \mathbf{u}\rangle) - \langle \boldsymbol{\beta}, [\mathbf{v}, \mathbf{u}]\rangle.$$

In other words $\mathbf{L_v}\boldsymbol{\beta} := \boldsymbol{\iota_v}\mathrm{d}\boldsymbol{\beta} + \mathrm{d}(\boldsymbol{\iota_v}\boldsymbol{\beta})$, a "magical" formula of *Cartan*. With this and the derivation property $\mathbf{L_v}$ may be extended to the whole tensor algebra $\mathscr{T}(\mathcal{M})$, and to the algebra of differential forms $\boldsymbol{\Lambda}(\mathcal{M})$.

The derivation property means that altogether we have for $\xi, \eta \in \mathscr{T}(\mathcal{M})$: $\mathbf{L_v}(\xi \otimes \eta) = (\mathbf{L_v}\xi) \otimes \eta + \xi \otimes (\mathbf{L_v}\eta)$, and for $\beta, \gamma \in \boldsymbol{\Lambda}(\mathcal{M})$ one has $\mathbf{L_v}(\beta \wedge \gamma) = (\mathbf{L_v}\beta) \wedge \gamma + \beta \wedge (\mathbf{L_v}\gamma)$.

11.3 Pull-Back and Push-Forward

Let $F : \mathcal{M} \to \mathcal{S}$ be a differentiable map. For functions, the **pull-back** $F^* : \mathscr{E}(\mathcal{S}) \to \mathscr{E}(\mathcal{M})$ of $\psi \in \mathscr{E}(\mathcal{S})$ is $F^*\psi := \psi \circ F \in \mathscr{E}(\mathcal{M})$, and the **push-forward** $F_* : \mathscr{E}(\mathcal{M}) \to \mathscr{E}(\mathcal{S})$ of $\varphi \in \mathscr{E}(\mathcal{M})$ is $F_*\varphi := \varphi \circ F^{-1} \in \mathscr{E}(\mathcal{S})$.

For vector fields, the **pull-back** $F^* : \mathscr{X}(\mathcal{S}) \to \mathscr{X}(\mathcal{M})$ of $\mathbf{w} \in \mathscr{X}(\mathcal{S})$ is $F^*\mathbf{w} := \mathbf{T}F^{-1} \circ \mathbf{w} \circ F \in \mathscr{X}(\mathcal{M})$, and the **push-forward** $F_* : \mathscr{X}(\mathcal{M}) \to \mathscr{X}(\mathcal{S})$ of $\mathbf{v} \in \mathscr{X}(\mathcal{M})$ is $F_*\mathbf{v} := \mathbf{T}F \circ \mathbf{v} \circ F^{-1} \in \mathscr{X}(\mathcal{S})$.

For differential forms, the **pull-back** $F^* : \mathscr{X}^*(\mathcal{S}) \to \mathscr{X}^*(\mathcal{M})$ of $\boldsymbol{\gamma} \in \mathscr{X}^*(\mathcal{S})$ is $F^*\boldsymbol{\gamma} := \mathbf{T}F^* \circ \boldsymbol{\gamma} \circ F \in \mathscr{X}^*(\mathcal{M})$, and the **push-forward** $F_* : \mathscr{X}^*(\mathcal{M}) \to \mathscr{X}^*(\mathcal{S})$ of $\boldsymbol{\beta} \in \mathscr{X}^*(\mathcal{M})$ is $F_*\boldsymbol{\beta} := \mathbf{T}F^{-*} \circ \boldsymbol{\beta} \circ F^{-1} \in \mathscr{X}^*(\mathcal{S})$.

11.4 Riemann Metric

A section $g : \mathcal{M} \to \Gamma^2(\mathscr{X}^*(\mathcal{M}))$ (this is a field of 2-fold covariant symmetric tensors, see section 8.2), such that at each $x \in \mathcal{M}$ the 2-form g_x is positive definite, is called a **Riemann metric**. Hence for $\boldsymbol{u}, \boldsymbol{v} \in \boldsymbol{T}_x\mathcal{M}$, the expression $g_x(\boldsymbol{u}, \boldsymbol{v})$ may be taken as *inner product* of the two vectors. A manifold $\mathcal{M}$ together with a *Riemannian metric* g is called a **Riemannian manifold** $(\mathcal{M}, g)$. For a function $\varphi \in \mathscr{E}(\mathcal{M})$, this allows to define the **gradient** as the *vector field* $\nabla\phi \in \mathscr{X}(\mathcal{M})$, such that for all $\mathbf{u} \in \mathscr{X}(\mathcal{M})$ it holds that $g(\nabla\varphi, \mathbf{u}) = \langle \mathrm{d}\varphi, \mathbf{u}\rangle = \mathbf{L_u}\,\varphi$. Observe that for the definition of this vector field one needs the Riemann metric, unlike the *co-vector field* $\mathrm{d}\varphi \in \mathscr{X}^*(\mathcal{M})$.

Given a C^1 curve $c : [0,1] \to \mathcal{M}$, the Riemann metric allows to define its length

$$L(c) := \int_0^1 \sqrt{g_t(c'(t), c'(t))}\,\mathrm{d}t,$$

where $c' \in \mathscr{X}(\mathcal{M})$ with $c'(t) \in \boldsymbol{T}_{c(t)}\mathcal{M}$ is the *derivative* of the curve at t.

A Riemann metric may be used to define a *real metric* τ on the manifold $\mathcal{M}$ via

$$\tau(x, y) := \inf\{L(c) |\, c \text{ is a piecewise } C^1 - \text{curve between } x \text{ and } y\}.$$

A curve realising the infimum is called a **geodetic curve**, and the metric is sometimes also called the **geodetic metric**.

A Riemann metric also allows to define the **Levi-Cívitá connection** or **covariant derivative** $\nabla_\mathbf{u}\mathbf{v}$, with $\nabla_\mathbf{u} \in \mathscr{L}(\mathscr{X}(\mathcal{M}))$, of a vector field $\mathbf{v} \in \mathscr{X}(\mathcal{M})$ with respect to another vector field $\mathbf{u} \in \mathscr{X}(\mathcal{M})$ —for functions $\varphi \in \mathscr{E}(\mathcal{M})$ it coincides with the normal directional derivative $\nabla_\mathbf{u}\varphi := \mathbf{L_u}\varphi(x) = g(\nabla\varphi, u) = \langle \mathrm{d}\varphi, \mathbf{u}\rangle$, hence here $\nabla_\mathbf{u} \in \mathscr{L}(\mathscr{E}(\mathcal{M}))$. It is a *derivation* and therefore satisfies $\nabla_\mathbf{u}(\varphi\mathbf{v}) = (\nabla_\mathbf{u}\varphi)\mathbf{v} + \varphi\nabla_\mathbf{u}\mathbf{v}$ for $\varphi \in \mathscr{E}(\mathcal{M})$. Similarly to the *Lie derivative* (see section 11.2), this allows the *covariant derivative* to be extended to the whole tensor algebra. It is also *torsion free* $\nabla_\mathbf{u}\mathbf{v} - \nabla_\mathbf{v}\mathbf{u} = [\mathbf{u}, \mathbf{v}] = \mathbf{L_u}\mathbf{v}$ (see section 11.2), and *compatible* with the Riemann metric $\langle \mathrm{d}g(\mathbf{v}, \mathbf{w}), \mathbf{u}\rangle = g(\nabla_\mathbf{u}\mathbf{v}, \mathbf{w}) + g(\mathbf{v}, \nabla_\mathbf{u}\mathbf{w}) = \boldsymbol{\iota}_\boldsymbol{u}\mathrm{d}g(\mathbf{v}, \mathbf{w}) = \mathbf{L_u}g(\mathbf{v}, \mathbf{w})$. It may be completely defined through the *Koszul* formula

$$\begin{aligned} 2\,g(\nabla_\mathbf{u}\mathbf{v}, \mathbf{w}) = \langle \mathrm{d}g(\mathbf{v}, \mathbf{w}), \mathbf{u}\rangle + \langle \mathrm{d}g(\mathbf{u}, \mathbf{w}), \mathbf{v}\rangle - \langle \mathrm{d}g(\mathbf{u}, \mathbf{v}), \mathbf{w}\rangle \\ - g(\mathbf{w}, [\mathbf{u}, \mathbf{v}]) + g(\mathbf{v}, [\mathbf{u}, \mathbf{w}]) + g(\mathbf{u}, [\mathbf{w}, \mathbf{v}]), \end{aligned}$$

as on the right hand side of the equation all terms have been defined previously and are well-known.

In contrast to the *Lie derivative* (see section 11.2), where no *Riemann metric* was needed, this is essential here.

The *Riemann metric* and its *Levi-Cívitá connection* or *covariant derivative* may be used to define the **Riemann curvature operator**

$$\forall \mathbf{u}, \mathbf{v} \in \mathscr{X}(\mathcal{M}): \ \boldsymbol{R}(\mathbf{u}, \mathbf{v}) := \nabla_{\mathbf{u}} \nabla_{\mathbf{v}} - \nabla_{\mathbf{v}} \nabla_{\mathbf{u}} - \nabla_{[\mathbf{u}, \mathbf{v}]},$$

such that for fixed $\mathbf{u}, \mathbf{v}$ this is a linear map $\mathscr{X}(\mathcal{M}) \ni \mathbf{w} \mapsto \boldsymbol{R}(\mathbf{u}, \mathbf{v})\mathbf{w} \in \mathscr{X}(\mathcal{M})$, i.e. $\boldsymbol{R} \in \mathscr{L}(\mathscr{X}(\mathcal{M})^2, \mathscr{L}(\mathscr{X}(\mathcal{M})))$. With this the **Riemann curvature tensor** $R \in \mathscr{T}_0^4(\mathscr{X}(\mathcal{M}))$ may be defined via

$$\forall \mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z} \in \mathscr{X}(\mathcal{M}): \ R(\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{z}) := g(\boldsymbol{R}(\mathbf{u}, \mathbf{v})\mathbf{w}, \mathbf{z}) \in \mathbb{R}.$$

12 Convex Analysis

If $\mathcal{X}$ is a vector space, and $f \in \mathscr{F}(\mathcal{X}, \overline{\mathbb{R}})$ a function(al), it is **convex** iff the **epigraph**

$$\text{epi } f := \{(x, \alpha) \in \mathcal{X} \times \mathbb{R}: \ f(x) \leq \alpha\}$$

is convex in $\mathcal{X} \times \mathbb{R}$. In this case the domain of f (cf. section 1.6) is

$$\text{dom f} = \{x \in \mathcal{X}: \ f(x) < \infty\}.$$

A convex function f is **proper**, iff dom $f \neq \emptyset$ and $\forall x \in \text{dom } f: \ f(x) > -\infty$.

For a minimisation $v = \inf_{x \in \mathcal{C}} f(x)$ with a convex function f over a convex set $\mathcal{C}$, it is advantageous to extend f by $f(x) + \psi_{\mathcal{C}}(x)$, as the same minimum (with less notation) is reached by $v = \inf_{x \in \mathcal{X}} \{f(x) + \psi_{\mathcal{C}}(x)\}$. The set of minimisers $\mathcal{M}^\flat(f) \subseteq \mathcal{X}$ is convex.

If f is convex, then $-f$ is **concave**. Note that affine functions are both. The sets $[f \leq \alpha] := \{x \in \mathcal{X}: \ f(x) \leq \alpha\}$ are called (lower) **sections**. The function f is **quasi-convex** iff $[f \leq \alpha]$ is convex for each α. The **convex hull** co f is the greatest convex function $\leq f$ (under the pointwise ordering in $\mathbb{R}$).

If f, g are convex functions, then the **infimal** or **inf** convolution is

$$f \Box g(x) := \inf_{y \in \mathcal{X}} \{f(x - y) + g(y)\}.$$

For a convex functional g on $\mathcal{Y}$ and $A \in \mathscr{L}(\mathcal{X}, \mathcal{Y})$, the **product from the right** gA on $\mathcal{X}$ is the composition $g(Ax)$. The **product from the left** is for a functional f on $\mathcal{X}$

$$\forall y \in \mathcal{Y}: \quad Af(y) := \inf_{Ax = y} f(x).$$

12.1 Subdifferential and Fenchel-Transform

Let $\mathcal{X}$ be a LCS, then a **subgradient** of a convex functional f at $x \in \mathcal{X}$ is any $x^* \in \mathcal{X}^*$ such that

$$\forall y \in \mathcal{X}: \quad f(x) + \langle x^*, y - x \rangle \leq f(y).$$

The set of all subgradients is the **subdifferential** of f, a closed convex subset denoted by $\partial f(x) \subset \mathcal{X}^*$. A convex function f is called **subdifferentiable** at $x \in \mathcal{X}$ iff $\partial f(x) \neq \emptyset$. For an extended convex f, it is clear that $x \in \mathcal{M}^\flat(f) \Rightarrow 0 \in \partial f(x)$. Note that if f is Gâteaux differentiable, $\partial f(x) = \{\delta f(x)\}$.

The function f is **lower semi-continuous** (lsc) iff $[f \leq \alpha]$ is closed for each α, or equivalently iff epi f is closed in $\mathcal{X} \times \mathbb{R}$. This means that for $x_j \to x$, one has $\liminf f(x_j) \leq f(x)$. The **lsc hull** lsc f is the greatest lsc function $\leq f$ (under the pointwise ordering in $\mathbb{R}$). The **closure** of a convex function f is

$$\mathrm{cl}\ f := \begin{cases} \mathrm{lsc}\ f, & \text{if } \forall x \in \mathcal{X}\ \mathrm{lsc}\ f(x) > -\infty \\ -\infty, & \text{otherwise.} \end{cases}$$

The **conjugate** (again a place for confusion) or **(Legendre-)Fenchel** transform of a functional f on $\mathcal{X}$ is

$$f^* : \mathcal{X}^* \to \overline{\mathbb{R}}, \quad f^*(x^*) = \sup_{x \in \mathcal{X}} \{\langle x^*, x \rangle - f(x)\}.$$

It is a convex function on $\mathcal{X}^*$ —as the supremum of affine functions $x^* \mapsto \langle x^*, x \rangle - f(x)$. Note the **Fenchel inequality:** $\forall x \in \mathcal{X}$, $x^* \in \mathcal{X}^*: \ f^*(x^*) + f(x) \geq \langle x^*, x \rangle$. Observe that on a reflexive space the **bi-conjugate** $f^{**} = \mathrm{cl\ co}\ f$. For a minimisation $v = \inf_{x \in \mathcal{X}} f(x)$, note that $-f^*(0) = v$.

For a proper, convex and lsc function f, one has

$$x^* \in \partial f(x) \ \Leftrightarrow \ x \in \partial f^*(x^*) \quad \text{and} \quad \langle x^*, x \rangle = f(x) + f^*(x^*).$$

Observe that the indicator function (section 1.6) of a convex set $\mathcal{A}$ is a convex function. For any positively homogeneous convex function of degree 1, one has $\partial f(x) = \{x^* | \ x^* \in \partial f(0) \wedge \langle x^*, x \rangle = f(x)\}$.

The **support function** of a set $\mathcal{A} \subseteq \mathcal{X}$ is $\psi^*_\mathcal{A}(x^*)$, it is a convex function, and $\psi^*_\mathcal{A}(x^*) = \sup_{x \in \mathcal{A}} \langle x^*, x \rangle$. If $0 \in \mathcal{A}$, then $\psi^*_\mathcal{A}$ is positively homogeneous of degree 1, and $\mathcal{A} = \partial \psi^*_\mathcal{A}(0)$. Observe that $\psi^*_\mathcal{A} = \psi^*_{\overline{\mathrm{co}}\,\mathcal{A}}$. For a set $\mathcal{A}$, the domain of the support functional dom $\psi^*_\mathcal{A} = \mathcal{A}^\infty \subseteq \mathcal{X}^*$ is a salient, pointed cone, the **barrier cone** of $\mathcal{A}$. It is convex iff $\mathcal{A}$ is so, and $0 \in \mathcal{A}: \ \mathcal{A}^\infty = (\mathcal{A}^\sharp)^\vee$. If $\mathcal{A}$ is a cone, $\psi^*_\mathcal{A} = \psi_{\mathcal{A}^-}$, the indicator of the **polar cone** $\mathcal{A}^-$.

The **gauge** or **Minkowski**-functional of a convex set $\mathcal{A}$ is a convex function defined by (where it is assumed that $0 \in \mathcal{A}$)

$$\mu_{\mathcal{A}}(x) := \inf\{\alpha > 0 | \, x \in \alpha\mathcal{A}\}.$$

In case $\mathcal{A}$ is weakly closed, absolutely convex, and absorbent (such sets are called **barrels**), the Minkowski-functional $\mu_{\mathcal{A}}$ is a seminorm on $\mathcal{X}$. If $\emptyset \neq \mathcal{A} \subseteq \mathcal{X}^*$ is weak* bounded and absolutely convex, $\psi^*_{\mathcal{A}}$ is a seminorm on $\mathcal{X}$. Observe that here $\psi^*_{\mathcal{A}} = \mu_{\mathcal{A}^o}$. And if $\mathcal{A}$ is a closed, convex set with $0 \in \mathcal{A}$, one has $\mathcal{A}^\infty = (\mathcal{A}^\wedge)^\ominus$ and $\mathcal{A} = \{x | \, \mu_{\mathcal{A}}(x) \leq 1\}$.

12.2 Cones and Polars

For a subset $\mathcal{C} \subseteq \mathcal{X}$ in a LCS, we have the

spanned or generated cone $\mathcal{C}^\vee = \bigcup_{\lambda \geq 0} \lambda\mathcal{C} \subseteq \mathcal{X}$.

recession cone $\mathcal{C}^\wedge = \bigcap_{\lambda \geq 0} \lambda\mathcal{C} \subseteq \mathcal{X}$.

(absolute) polar $\mathcal{C}^o = \{x^* | \forall x \in \mathcal{C} : \, |\langle x^*, x\rangle| \leq 1\} \subseteq \mathcal{X}^*$.

lower polar $\mathcal{C}^\flat = \{x^* | \forall x \in \mathcal{C} : \, \langle x^*, x\rangle \geq -1\} \subseteq \mathcal{X}^*$.

upper polar $\mathcal{C}^\sharp = \{x^* | \forall x \in \mathcal{C} : \, \langle x^*, x\rangle \leq 1\} \subseteq \mathcal{X}^*$.

negative polar cone $\mathcal{C}^\ominus = \{x^* | \forall x \in \mathcal{C} : \, \langle x^*, x\rangle \leq 0\} \subseteq \mathcal{X}^*$.

positive polar cone $\mathcal{C}^\oplus = \{x^* | \forall x \in \mathcal{C} : \, \langle x^*, x\rangle \geq 0\} \subseteq \mathcal{X}^*$.

annihilator cone $\mathcal{C}^\perp = \{x^* | \forall x \in \mathcal{C} : \, \langle x^*, x\rangle = 0\} \subseteq \mathcal{X}^*$.

barrier cone $\mathcal{C}^\infty = \{x^* | \psi^*_{\mathcal{C}}(x^*) = \sup_{x \in \mathcal{C}} \langle x^*, x\rangle < \infty\} = \operatorname{dom} \psi^*_{\mathcal{C}} \subseteq \mathcal{X}^*$.

Observe that $\mathcal{C}^\wedge \subseteq \mathcal{C}^\vee$. Also note that $\mathcal{C}^o \subseteq \mathcal{C}^\sharp \subseteq \mathcal{C}^\infty$, as well as $\mathcal{C}^\ominus \subseteq \mathcal{C}^\sharp$, and $\mathcal{C}^\perp \subseteq \mathcal{C}^o$, and $\mathcal{C}^\perp = \mathcal{C}^\ominus \cap \mathcal{C}^+$.

12.3 Perturbations, Lagrangians, and Hamiltonians

For a **primal** minimisation problem $v = \inf_{x \in \mathcal{X}} f(x)$ consider **perturbations** $F : \mathcal{X} \times \mathcal{U} \to \overline{\mathbb{R}}$ —where $\mathcal{U}$ is another LCS—such that $\forall x \in \mathcal{X} : \;\; f(x) = F(x, 0)$. Consider the **perturbed** minimisation problem $v(u) = \inf_{x \in \mathcal{X}} F(x, u)$, such that $v = v(0)$. If F is convex, so is the **minimal value** function $v : \mathcal{U} \to \overline{\mathbb{R}}$.

With the perturbations F, a **Lagrangian** may be defined as

$$\ell(x, u^*) := \inf_{u \in \mathcal{U}} \{F(x, u) + \langle u^*, u\rangle\}.$$

Note that $f(x) = \sup_{u^* \in \mathcal{U}^*} \ell(x, u^*)$ under appropriate conditions. The dual problem is defined via $g(u^*) = \inf_{x \in \mathcal{X}} \ell(x, u^*)$; where the function g is concave, and

$$G(u^*, x^*) := \inf_{x \in \mathcal{X}} \{\ell(x, u^*) + \langle x^*, x\rangle\},$$

so that $g(u^*) = G(u^*, 0)$. Note that $g(u^*) = \inf_{x \in \mathcal{X}} \ell(x, u^*)$ under appropriate conditions. With g and G associate the **dual problem** $\gamma = \sup_{u^* \in \mathcal{U}^*} g(u^*)$, and the optimal value function $\gamma(x^*) = \sup_{u^* \in \mathcal{U}^*} G(u^*, x^*)$. Solutions to the dual problem $\bar{u}^* \in \mathcal{U}^*$ are **Lagrange multipliers** of the primal problem.

It is clear that

$$\forall x \in \mathcal{X}, u^* \in \mathcal{U}^* : \; f(x) = \sup_{v^* \in \mathcal{U}^*} \ell(x, v^*) \geq \ell(x, u^*) \geq \inf_{y \in \mathcal{X}} \ell(y, u^*) = g(u^*),$$

and therefore

$$\upsilon = \inf_{x \in \mathcal{X}} f(x) = \inf_{x \in \mathcal{X}} \sup_{u^* \in \mathcal{U}^*} \ell(x, u^*) \geq \sup_{u^* \in \mathcal{U}^*} \inf_{x \in \mathcal{X}} \ell(x, u^*) = \sup_{u^* \in \mathcal{U}^*} g(u^*) = \gamma.$$

If equality $\upsilon = \gamma$ holds, this is called the **saddle-value**, and a point $(\bar{x}, \bar{u}^*) \in \mathcal{X} \times \mathcal{U}^*$ such that

$$\forall x \in \mathcal{X}, u^* \in \mathcal{U}^* : \; \ell(x, \bar{u}^*) \geq \ell(\bar{x}, \bar{u}^*) \geq \ell(\bar{x}, u^*)$$

is a **saddle-point**. Observe that $(\bar{x}, \bar{u}^*)$ is a saddle-point iff $\bar{x}$ solves the primal minimisation problem, i.e. $\upsilon = f(\bar{x})$, and $\bar{u}^*$ solves the dual problem, $\gamma = g(\bar{u}^*)$, and one has $\bar{x} \in \partial\gamma(0)$ and $\bar{u}^* \in \partial\upsilon(0)$.

For the special primal minimisation problem $\upsilon = \inf_{x \in \mathcal{X}} F(x, Ax)$, where $A \in \mathscr{L}(\mathcal{X}, \mathcal{U})$, define the **Hamiltonian**

$$H(x, u^*) := \sup_{u \in \mathcal{U}} \{\langle u^*, u\rangle - F(x, u)\}.$$

If $\bar{x} \in \mathcal{X}$ is a solution to $\upsilon = \inf_{x \in \mathcal{X}} F(x, Ax)$ and $\bar{u}^* \in \mathcal{U}^*$ is a *Lagrange multiplier*, one has the **Euler-Lagrange** relations

$$(-A^*\bar{u}^*, \bar{u}^*)^T \in \partial F(\bar{x}, A\bar{x}).$$

This is equivalent to the **Hamiltonian** relations (with *partial* subdifferentials)

$$A\bar{x} \in \partial_{u^*} H(\bar{x}, \bar{u}^*) \text{ and } A^*\bar{u}^* \in \partial_x H(\bar{x}, \bar{u}^*).$$

13 Stochastics

A random variable (RV) $X \in L_0(\Omega, \mathfrak{A}, \mathbb{P})$ is a measurable map on a **probability space**, i.e. a measurable space with a finite measure $\mathbb{P}$, the **probability measure**, normed such that $\Pr(\Omega) := \mathbb{P}(\Omega) = 1$ (total probability). Associated to a random variable is its **distribution measure**, the *push-forward* of $\mathbb{P}$ (also called the **law** of X) $\mu_X := X_*\mathbb{P}$ on $(\mathbb{R}, \mathfrak{B}_\mathbb{R})$ —the real numbers with the *Borel* sets, given for $A \in \mathfrak{B}_\mathbb{R}$ by $\Pr_X(A) = \mu_X(A) := X_*\mathbb{P}(A) = \mathbb{P}(X^{-1}(A))$. Often subsets of Ω are defined by $X^{-1}(A)$, where $A \subseteq \mathbb{R}$ (e.g. $\{\omega | X(\omega) \leq c\}$ for $c \in \mathbb{R}$). Then $\mathbb{P}(X^{-1}(A)) = \mu_X(A)$, and this is often abbreviated to e.g. $\Pr(X \leq c) = \mathbb{P}(X \leq c) := \mu_X(]-\infty, c]) = \mathbb{P}(X^{-1}(]-\infty, c]))$.

13.1 Distribution Function and Density

The **distribution function** of X on $\mathbb{R}$ is given by $\forall x \in \mathbb{R}: \quad F_X(x) := \Pr_X(]-\infty, x]) = \mu_X(]-\infty, x]) = \mathbb{P}(X \leq x) = \Pr_X(X \leq x)$. Of course F_X is non-negative and non-decreasing, with obviously $\lim_{x\to-\infty} F_X(x) = 0$ and $\lim_{x\to\infty} F_X(x) = 1$. With this function, one may write for any $\varphi \in C_b(\mathbb{R})$

$$\langle \mu_X, \varphi \rangle = \langle X_*\mathbb{P}, \varphi \rangle = \int_\mathbb{R} \varphi \, \mathrm{d}\mu_X = \int_\mathbb{R} \varphi(x) \, \mu_X(\mathrm{d}x) = \int_\mathbb{R} \varphi(x) \, \mathrm{d}F_X(x),$$

where the last integral is a **Lebesgue-Stieltjes** integral. Observe that naturally $\langle \mu_X, 1 \rangle = \int_\mathbb{R} \mathrm{d}\mu_X = 1$. Obviously $\mu_X = X_*\mathbb{P}$ is also a continuous functional on $\mathscr{D}(\mathbb{R}) = C_c^\infty(\mathbb{R})$.

The derivative in the sense of distributions of F_X (or μ_X) is the **probability density** $f_X := \frac{\mathrm{d}}{\mathrm{d}x} F_X$, and $0 \leq f_X$ in the order induced on $\mathscr{D}'(\mathbb{R})$ through the normal pointwise order on $\mathscr{D}(\mathbb{R})$. Sometimes one only speaks of *probability density* in case $f_X \in L_1(\mathbb{R})$ is a *regular* distribution (in that case μ_X is *absolutely continuous*—see section 3—w.r.t. *Lebesgue measure* on $\mathbb{R}$, and $\|f_X\|_{L_1} = 1$), although most things are simpler with the more general definition.

13.2 Moments and Characteristic Function

In case $X \in L_1(\Omega, \mathbb{P})$, the number $\mathbb{E}(X) := \int_\Omega X \, \mathrm{d}\mathbb{P}$ is the **expectation, mean** or **expected value** of X, also denoted by $\langle X \rangle$ or $\bar{X}$. This is also given by $\mathbb{E}(X) = \int_\mathbb{R} x \, \mu_X(\mathrm{d}x) = \int_\mathbb{R} x \, \mathrm{d}F_X(x) = \int_\mathbb{R} x \, f_X(x) \, \mathrm{d}x$. For $X \in L_2(\Omega, \mathbb{P}) \subset L_1(\Omega, \mathbb{P})$, the **variance** of X is the expected value of $(X - \bar{X})^2$. More generally, for $k \in \mathbb{N}_0$, the ***k*-th moment** is $\mu_k := \mathbb{E}(X^k)$, and the ***k*-th central moment** (with $\tilde{X} := X - \bar{X}$ the zero-mean random part) is $\tilde{\mu}_k := \mathbb{E}(\tilde{X}^k)$.

The *Fourier transform* of a random variable (RV) X —or rather its *distribution measure*—is its **characteristic function** $\varphi_X(y) := \hat{X}(y) := \mathbb{E}(\exp(\mathrm{i}yX)) = \int_{\mathbb{R}} \exp(\mathrm{i}yx)\,\mu_X(\mathrm{d}x)$. Note that $\mu_k = \mathrm{i}^k \frac{\mathrm{d}^k}{\mathrm{d}y^k}\varphi_X(0)$. While this function is defined for all $y \in \mathbb{R}$, the *Laplace transform* of X —or rather its *distribution measure*—is the **moment generating function** $\psi_X(z) := \mathbb{E}(\exp(zX)) = \int_{\mathbb{R}} \exp(zx)\,\mu_X(\mathrm{d}x)$, and it may not be defined for all $z \in \mathbb{C}$, but certainly near the origin, and one has $\mu_k = \frac{\mathrm{d}^k}{\mathrm{d}z^k}\psi_X(0)$. Hence, where defined, $\psi_X(z) = \sum_{k=0}^{\infty} \mu_k \frac{z^k}{k!}$. The **cumulant function** is $\kappa_X(z) := \ln(\psi_X(z)) = \sum_{k=1}^{\infty} \kappa_k \frac{z^k}{k!}$, and the coefficients $\kappa_k = \frac{\mathrm{d}^k}{\mathrm{d}z^k}\kappa_X(0)$ are the k**-th cumulants** of the RV X.

Closely related is the Θ- or **Hermite-Laplace transform** $\vartheta_X(z) := e^{z^2/2}\psi_X(z) = \sum_{k=0}^{\infty} \nu_k \frac{z^k}{k!}$, and the coefficients $\nu_k = \frac{\mathrm{d}^k}{\mathrm{d}z^k}\vartheta_X(0)$ are the k**-th quasi-moments** of X. If the RV $Y = f(X) \in L_2(\mu_X)$ is a function of a Gaussian RV X with zero mean $\mathbb{E}(X) = 0$ and unit variance $\mathbb{E}(X^2) = 1$, the quasi-moments are generated from the k-th *Hermite polynomial* $h_k(t) := (-1)^k e^{t^2/2} \frac{\mathrm{d}^k}{\mathrm{d}t^k} e^{-t^2/2}$ via $\nu_k = \mathbb{E}(Y h_k(X)) = \mathbb{E}(f(X)h_k(X)) = \mathbb{E}(\mathrm{D}^k f(X))$. The RV $Y = f(X)$ may then be written as a series of polynomials of the Gaussian RV X as $Y = \sum_{k=0}^{\infty} \nu_k \frac{h_k(X)}{k!}$, convergent in $L_2(\mu_X)$ —the scaled Hermite polynomials $\{h_k(t)/\sqrt{k!}\}_{k=0}^{\infty}$ are a CONS in $L_2(\mu_X)$. This is *Wiener's* **polynomial chaos** expansion (PCE).

13.3 Vector Valued Random Variables

For a LCS $\mathcal{X}$ with the *Borel algebra* $\mathfrak{B}_{\mathcal{X}}$ —see section 3, a measurable mapping $X \in L_0(\Omega; \mathcal{X}) \subseteq \mathscr{F}(\Omega, \mathcal{X})$ is called a $\mathcal{X}$-valued random variable. The measure $\mu_X := X_*\mathbb{P}$ induced on $(\mathcal{X}, \mathfrak{B}_{\mathcal{X}})$ via $\mu_X(\mathcal{A}) := \mathbb{P}(X^{-1}(\mathcal{A})$ for $\mathcal{A} \in \mathfrak{B}_{\mathcal{X}}$ is again called the **distribution measure** of X.

In case $X \in L_1(\Omega, \mathbb{P}; \mathcal{X})$, the vector $\mathbb{E}(X) := \int_\Omega X\,\mathrm{d}\mathbb{P}$ is the **expectation** or **expected value** of X, also denoted by $\langle X \rangle$ or $\bar{X}$. If $X, Y \in L_2(\Omega, \mathbb{P}; \mathcal{X})$, the **correlation** is the bilinear form $\mathbb{E}(X \otimes Y)$, and the **covariance** the bilinear form $\mathrm{Cov}_X := \mathbb{E}((X - \bar{X}) \otimes (Y - \bar{Y}))$; and closely related is the **covariance operator** $C_X : \mathcal{X}^* \ni x^* \mapsto \mathrm{Cov}_X(x^*, \cdot) \in (\mathcal{X}^*)'$. More generally, for $X \in L_k(\Omega, \mathbb{P}; \mathcal{X})$ and $k \in \mathbb{N}_0$, the k**-th moment** is $\mu_k := \mathbb{E}(X^{\otimes k}) \in \mathcal{X}^{\vee k}$, and the k**-th central moment** (with $\tilde{X} := X - \bar{X}$ the zero-mean random part) is $\tilde{\mu}_k := \mathbb{E}(\tilde{X}^{\otimes k}) \in \mathcal{X}^{\otimes k}$.

Again the *Fourier transform* of a random variable (RV) X —or rather its *distribution measure*—is a function defined on the dual $\mathcal{X}^*$, its **characteristic function** $\varphi_X(x^*) := \mathbb{E}(\exp(\mathrm{i}\langle x^*, X\rangle)) = \int_{\mathcal{X}} \exp(\mathrm{i}\langle x^*, x\rangle)\,\mu_X(\mathrm{d}x) \in \mathscr{F}(\mathcal{X}^*, \mathbb{C})$. Note that $\mu_k = \mathrm{i}^k \mathrm{D}^k \varphi_X(0)$. While this function is defined for all $x^* \in \mathcal{X}^*$, the *Laplace transform* of X —or rather its *distribution measure*—is the **moment generating function** $\psi_X(z) := \mathbb{E}(\exp(\langle z, X\rangle)) = \int_{\mathcal{X}} \exp(\mathrm{i}\langle z, x\rangle)\, \mu_X(\mathrm{d}x) \in \mathscr{F}(\mathcal{X}_c^*, \mathbb{C})$, and it is not defined for all $z \in \mathcal{X}_c^* := \mathcal{X} \oplus \mathrm{i}\mathcal{X}$ in the *complexification* of $\mathcal{X}$, but one has $\mu_k = \mathrm{D}^k\psi_X(0) \in \mathcal{X}^{\vee k}$. Hence, where defined, $\psi_X(z) = \sum_{k=0}^{\infty} \frac{1}{k!}\mu_k z^{\vee k}$. The **cumulant function** is $\kappa_X(z) := \ln(\psi_X(z)) = \sum_{k=1}^{\infty} \frac{1}{k!}\kappa_k z^{\vee k}$, and the coefficients $\kappa_k = \mathrm{D}^k\kappa_X(0) \in \mathcal{X}^{\vee k}$ are the **k-th cumulants** of the RV X.

Closely related is the Θ- or **Hermite-Laplace transform** $\vartheta_X(z) := e^{\langle z,z\rangle/2}\psi_X(z) = \sum_{k=0}^{\infty} \frac{1}{k!}\nu_k z^{\vee k}$, and the coefficients $\nu_k = \mathrm{D}^k\vartheta_X(0) \in \mathcal{X}^{\vee k}$ are the **k-th quasi-moments** of X.

13.4 Integration Algebra

Another view [23], which treats RVs more as primary objects and does not demand from the outset to define a measure space with measurable functions, is that the RVs are elements of an associative algebra, or to be more precise, a B*-algebra $\mathcal{B}$ over $\mathbb{C}$. Such an algebra is as vector space a *Banach* space (see section 4.1), the algebra multiplication is continuous (it satisfies $\forall x, y \in \mathcal{B} : \|xy\| \leq \|x\|\|y\|$), and has an algebra-anti-automorphism, which is an *involution* (see section 1.6), denoted by $* : x \mapsto x^*$, such that (the 'anti'-linearity) $\forall \alpha \in \mathbb{C} : \ *(\alpha x) = \bar{\alpha} * (x) = \bar{\alpha}x^*$. The B*-condition is that $\|x\| = \|x^*\|$ (C*-algebras in addition satisfy $\|x^*x\| = \|x\|^2$). For simplicity, assume that the algebra has an identity e, where necessarily one has $e^* = e$. Elements y which may be written as $y = x^*x$ are called **positive**, and if $x^* = x$, the element x is called **self-adjoint**. Positive elements are self-adjoint, and the set of positive elements is a convex cone, defining a partial order relation (see section 1.5) $0 \leq x$. For positive elements $0 \leq x$ and any $k \in \mathbb{N}$ there is a positive y such that $y^k = x$ (the positive k-th root).

Of course $\mathbb{C}$ with complex conjugation as involution is both a B*- and a C*-algebra, as is $\mathscr{F}(\mathcal{A}, \mathbb{C})$ ($\mathcal{A}$ any set), and its sub-algebras like $\mathscr{B}(\mathcal{A}, \mathbb{C})$, or $L_\infty(\mathcal{A}, \mathbb{C})$, $\mathscr{B}_b(\mathcal{A}, \mathbb{C})$, and $C(\mathcal{A}, \mathbb{C})$. These examples are all also *commutative*, a non-commutative example is $\mathbb{C}^{(n\times n)}$ and its sub-algebras, with $*(\boldsymbol{A}) := \boldsymbol{A}^*$ for $\boldsymbol{A} \in \mathbb{C}^{(n\times n)}$ (see section 1.9), as is the algebra of operators on a complex Hilbert space $\mathscr{L}(\mathcal{H})$, and its sub-algebras like the compact operators $\mathscr{L}_0(\mathcal{H})$.

One assumes that on this algebra a positive linear functional $\mathbb{E} \in \mathscr{L}(\mathcal{B}, \mathbb{C})$ is given (an **abstract integral**), such that

- $\mathbb{E}(x^*) = \overline{\mathbb{E}(x)}$, i.e. $\mathbb{E}$ commutes with involution.

- $\mathbb{E}(x) \geq 0$ if $0 \leq x$, the positivity condition, i.e. $\mathbb{E}$ is isotone. By passing to a factor algebra if necessary, one may assume that $\mathbb{E}(x) = 0$ iff $x = 0$.

- $\mathbb{E}(xy) = \mathbb{E}(yx)$.
- $|\mathbb{E}(x^*yx)| \leq c(y)\mathbb{E}(x^*x)$, where $\mathbb{R} \ni c(y) > 0$ depends only on y.

Most obvious examples are for a topological measure space the already mentioned B*-algebras (with pointwise multiplication) $L_\infty(\mathcal{A}, \mathbb{C})$, $\mathscr{B}_b(\mathcal{A}, \mathbb{C})$, and $C(\mathcal{A}, \mathbb{C})$, or $L_1(\mathcal{A}, \mathbb{C})$ (with convolution the algebra multiplication), and $\mathbb{E}$ the usual integral. If $\mathscr{A} \subset \mathscr{L}(\mathcal{H})$ is a commutative B*-sub-algebra and $x \in \mathcal{H}$ any vector, the functional may be defined for $A \in \mathscr{A}$ via $\mathbb{E}(A) := \langle Ax|x\rangle_{\mathcal{H}}$ and the *adjoint* as *-involution. Another example for $\boldsymbol{A} \in \mathbb{C}^{(n\times n)}$ (with the conjugate transpose as *-involution) is $\mathbb{E}(\boldsymbol{A}) := \operatorname{tr} \boldsymbol{A}$, which may also be taken for a *nuclear operator* $A \in \mathscr{L}_1(\mathcal{H})$ (see section 4.6.1) with $\mathbb{E}(A) := \operatorname{tr} A$.

As via the *Gelfand*-homomorphism a commutative B*-algebra is isomorphic to the continuous functions on its *spectral space* or *spectrum*—the subset $\mathcal{S} \subset \mathcal{B}^* \setminus \{0\}$ of the dual which are also **-algebra homomorphisms* into $\mathbb{C}$—a functional like $\mathbb{E}$ defines an element of the dual, i.e. a *Baire* measure (see section 3). This measure may then be used to define an integral on $\mathcal{S}$, connecting integration algebras with the usual measure theoretic approach. As one may set $\mathbb{E}(e) = 1$, the measure then is a *probability measure*.

For a *real* algebra, the functional $\mathbb{E}$ is *real-valued*. In case $\mathcal{B} = C(\mathcal{A}, \mathbb{R})$, the *-involution may be taken as the identity, in case $\mathcal{B} = \mathbb{R}^{(n\times n)}$ the *-involution is the *transpose* $\boldsymbol{A} \mapsto \boldsymbol{A}^T$, and for $A \in \mathscr{L}(\mathcal{H})$ or $A \in \mathscr{L}_1(\mathcal{H})$ it could be the *adjoint*.

Via $\langle x|y\rangle_{\mathcal{B}} := \mathbb{E}(x\, y^*)$ an inner product may be defined on $\mathcal{B}$, making it into a *pre-Hilbert space*. Multiplication by a fixed element $a \in \mathcal{B}$ such that $L_a : x \mapsto ax$ is then a *representation* of $\mathcal{B}$ in an algebra of continuous operators on a Hilbert space.

This approach may be in many cases simpler and more natural, in particular when considering random variables which are not *real-* or *complex-valued*, i.e. random processes and fields.

13.5 Convergence

Convergence of a sequence of RVs $X_n(\omega)$ on a probability space $(\Omega, \mathfrak{A}, \mathbb{P})$ towards a RV $X(\omega)$:

uniformly : Convergence in the strong topology of $\mathscr{B}_b(\Omega, \mathfrak{A})$ with the $\|\cdot\|_\infty$-norm.

almost uniformly : For any $\epsilon > 0\, \exists A \in \mathfrak{A}$ with $\mathbb{P}(A) < \epsilon$ such that the sequence converges in the strong topology of $\mathscr{B}_b(\Omega \setminus A, \mathfrak{A})$ with the $\|\cdot\|_\infty$-norm.

essentially uniformly : Convergence of the sequence in the strong topology of $L_\infty(\Omega, \mathfrak{A}, \mathbb{P})$ with the $\|\cdot\|_\infty$-norm. A special case of convergence in the p-th *mean* with $p = \infty$.

pointwise (does not make much sense as X_n may be changed on sets of measure zero): $\forall \omega \in \Omega : \; X_n(\omega) \to X(\omega)$. This is convergence in the *simple* or *product* topology on $\mathscr{F}(\Omega, \mathbb{R}) = \mathbb{R}^\Omega$. Denoted as $X_n \overset{\mathrm{s}}{\to} X$. Coincides with order convergence in $\mathscr{B}_b(\Omega, \mathfrak{A})$

pointwise a.s. or a.e. : $\exists N \in \mathfrak{A}, \mathbb{P}(N) = 0, \; \forall \omega \in \Omega \setminus N : \; X_n(\omega) \to X(\omega)$. This is convergence in the *final* topology on $L_0(\Omega, \mathfrak{A})$ inherited from the *simple* or *product* topology on $\mathscr{F}(\Omega, \mathbb{R}) = \mathbb{R}^\Omega$. Denoted as $X_n \overset{\mathrm{a.s.}}{\to} X$. Coincides with order convergence in $L_0(\Omega, \mathfrak{A})$.

in p-th mean : in the strong topology of $L_p(\Omega, \mathfrak{A}, \mathbb{P})$, usually with $p \geq 1$. May be simply denoted as $X_n \to X$ as it is the *strong convergence* of RVs, or also by $X_n \overset{\mathrm{m.p.}}{\to} X$. Especially with $p = 2$, this is called convergence in **mean square**, and is denoted by $X_n \overset{\mathrm{m.s.}}{\to} X = \mathrm{l.i.m.} X_n$, meaning **limit in mean** square.

in probability (in measure): in the metric topology of $L_0(\Omega, \mathfrak{A}, \mathbb{P})$, see section 3. Sometimes denoted as $X_n \overset{\mathbb{P}}{\to} X$.

weakly in L_p in a functional analytic sense: in the weak topology of some $L_p(\Omega, \mathfrak{A}, \mathbb{P})$. Denoted as $X_n \rightharpoonup X$.

weakly in measure : convergence of distribution measure on $C_b(\Omega)$.

weak* in measure/ in distribution : convergence of distribution measure on $C_0(\Omega)$. This is the *weak** topology on the measure space. Denoted as $X_n \overset{*}{\rightharpoonup} X$, or as $X_n \overset{\mathrm{d}}{\to} X$.

vaguely in measure : convergence of distribution measure on $C_c(\Omega)$. May be denoted as $X_n \overset{\mathrm{v}}{\to} X$.

14 Available Fonts

First look at different fonts in text mode, c.f. section 14.1. The commands have been abbreviated in the `nfontdef.sty` file. After that consider special symbols section 14.2 and section 14.3, and finally fonts in math-mode, c.f. section 14.4.

14.1 Fonts in Text

- *Emphasised* text is typed like this: `\emph{Emphasised}`.
  This is emphasised *text*, which is written as
  `{\em This is {\em emphasised} text}`.

- **Bold** text is typed like this: `\tbf{Bold}`.
  This is bold text, which is written as
  `{\bff This is bold text}`.

- *Italic* text is typed like this: `\tit{Italic}`.
  This is italic text, which is written as
  `{\itf This is italic text}`.

- *Slanted* text is typed like this: `\tsl{Slanted}`.
  This is slanted text, which is written as
  `{\slf This is slanted text}`.

- Sans Serif text is typed like this: `\tsf{Sans Serif}`.
  This is sans serif text, which is written as
  `{\sff This is sans serif text}`.

- `Typewriter` text is typed like this: `\ttt{Typewriter}`.
  `This is typewriter text` which is written as
  `{\ttf This is typewriter text}`.

- SMALL CAPS text is typed like this: `\tsc{Small Caps}`.
  THIS IS SMALL CAPS TEXT, which is written as
  `{\scf This is small caps text}`.

14.2 Non-ASCII Characters

All the TEX-Ways for “Umlaute” naturally work too: “äÄ öÖ üÜ ß éè âÂ ÈÉ çÇ ñÑ ÿë ïÏ šŠ čČ ćĆ ŏ ő ọ o̱ o͡o ō ȯ ¶§? ¿ ! ¡ œŒæÆåÅ©... ”.

14.3 Special Symbols

$ & % # _ { } [ ] are easy to produce.

14.4 Fonts in Math

Normal math font: $abcdefl\ell ABCDEF$, which is usually the same as

Italic math font: *abcdefABCDEF*, or $y = Ax$.

Bold math font: $\boldsymbol{abcdefl\ell ABCDEF}$, or $\boldsymbol{y} = \boldsymbol{Ax}$.

SansSerif math font: $\mathsf{abcdeflABCDEF}$, or $\mathsf{y} = \mathsf{Ax}$.

Bold SansSerif math font: $\boldsymbol{\mathsf{abcdeflABCDEF}}$, or $\boldsymbol{\mathsf{y}} = \boldsymbol{\mathsf{Ax}}$.

Roman math font: $\mathrm{abcdeflABCDEF}$

Bold Roman math font: $\mathbf{abcdeflABCDEF}$, or $\mathbf{y} = \mathbf{Ax}$.

Typewriter math font: $\mathtt{abcdeflABCDEF}$

And some fancier fonts:

EF *Euler Fraktur* math font: $\mathfrak{abcdeflABCDEFLUVWXYZ}$, or $\mathfrak{l} \in \mathfrak{A}$.

Cal *Calligraphic* math font: $\mathcal{ABCDEFLUVWXYZ}$

BBD *Black Board Double* math font: $\mathbb{ABCDEFLNZQRCXY}$

Scr *Script* math font: $\mathscr{ABCDEFLUVWXYZ}$

References

[1] ISO 31-11: Mathematical Notation, `http://en.wikipedia.org/wiki/ISO_31-11`

[2] R. Abraham, J. E. Marsden, T. Ratiu: *Manifolds, Tensor Analysis, and Applications*. Addison-Wesley, Reading, MA, 1983.

[3] N. I. Achieser, I. M. Glasmann: *Theorie der linearen Operatoren im Hilbert-Raum*. Akademie-Verlag, Berlin, 1968.

[4] D. Adams: *The Hitchhiker's Guide to the Galaxy*. Pocket Books, New York, NY, 1979.

[5] Ch. D. Aliprantis, K. C. Border: *Infinte Dimensional Analysis. A Hitchhiker's Guide*. Springer, Berlin, 1994.

[6] J. P. Aubin: *Mathematical Methods of Game and Economic Theory*. North-Holland, Amsterdam, 1979.

[7] Y. Choquet-Bruhat, C. de Witt-Morette, M. Dillard-Bleick: *Analysis, Manifolds, and Physics*. North-Holland, Amsterdam, 1981.

[8] J. B. Conway: *A Course in Functional Analysis*. Springer, Berlin, 1997.

[9] J. Dieudonné: *Grundzüge der modernen Analysis*. Vieweg, Braunschweig, 1971.

[10] I. Ekeland, R. Temam: *Convex Analysis and Variational Problems*. North-Holland, Amsterdam, 1976.

[11] G. Grätzer: *Math into LaTeX*. Birkhäuser, Basel, 1996.

[12] W. H. Greub: *Linear Algebra*. Springer, Berlin, 1967.

[13] W. H. Greub: *Multilinear Algebra*. Springer, Berlin, 1967.

[14] H. Heuser: *Funktionalanalysis*. Teubner, Stuttgart, 1992.

[15] F. Hirzebruch, W. Scharlau: *Einführung in die Funktionalanalysis*. Bibliographisches Institut, Mannheim, 1971.

[16] R. B. Holmes: *Geometric Functional Analysis and its Applications*. Springer, Berlin, 1975.

[17] S. Janson: *Gaussian Hilbert Spaces*. Cambridge University Press, Cambridge, 1997.

[18] H. Jarchow: *Locally Convex Spaces*. Teubner, Stuttgart, 1981.

[19] P. Krée, Chr. Soize: *Mathematics of Random Phenomena*. D. Reidel, Dordrecht, 1986.

[20] P. D. Lax: *Linear Algebra*. John Wiley & Sons, Chichester, 1996.

[21] R. T. Rockafellar: *Conjugate Duality and Optimization*. SIAM, Philadelphia, 1974.

[22] E. Schechter: *Handbook of Analysis and Its Foundations*. Academic Press, San Diego, 1997.

[23] I. E. Segal, R. A. Kunze: *Integrals and Operators*. Springer, Berlin, 1978.

[24] D. Werner: *Funktionalanalysis*. Springer, Berlin, 2005.

[25] K. Yosida: *Functional Analysis*. Springer, Berlin, 1995.

Technische Universität Braunschweig
Informatik-Berichte ab Nr. 2008-06

2008-06	M. Broy, M. V. Cengarle, H. Grönniger B. Rumpe	Modular Description of a Comprehensive Semantics Model for the UML (Version 2.0)
2008-07	C. Basarke, C. Berger, K. Berger, K. Cornelsen, M. Doering J. Effertz, T. Form, T. Gülke, F. Graefe, P. Hecker, K. Homeier F. Klose, C. Lipski, M. Magnor, J. Morgenroth, T. Nothdurft, S. Ohl, F. Rauskolb, B. Rumpe, W. Schumacher, J. Wille, L. Wolf	2007 DARPA Urban Challenge Team CarOLO - Technical Paper
2008-08	B. Rosic	A Review of the Computational Stochastic Elastoplasticity
2008-09	B. N. Khoromskij, A. Litvinenko, H. G. Matthies	Application of Hierarchical Matrices for Computing the Karhunen-Loève Expansion
2008-10	M. V. Cengarle, H. Grönniger B. Rumpe	System Model Semantics of Statecharts
2009-01	H. Giese, M. Huhn, U. Nickel, B. Schätz (Herausgeber)	Tagungsband des Dagstuhl-Workshops MBEES: Modellbasierte Entwicklung eingebetteter Systeme V
2009-02	D. Jürgens	Survey on Software Engineering for Scientific Applications: Reuseable Software, Grid Computing and Application
2009-03	O. Pajonk	Overview of System Identification with Focus on Inverse Modeling
2009-04	B. Sun, M. Lochau, P. Huhn, U. Goltz	Parameter Optimization of an Engine Control Unit using Genetic Algorithms

2009-05	A. Rausch, U. Goltz, G. Engels, M. Goedicke, R. Reussner	LaZuSo 2009: 1. Workshop für langlebige und zukunftsfähige Softwaresysteme 2009
2009-06	T. Müller, M. Lochau, S. Detering, F. Saust, H. Garbers, L. Märtin, T. Form, U. Goltz	Umsetzung eines modellbasierten durchgängigen Enwicklungsprozesses für AUTOSAR-Systeme mit integrierter Qualitätssicherung
2009-07	M. Huhn, C. Knieke	Semantic Foundation and Validation of Live Activity Diagrams
2010-01	A. Litvinenko, H. G. Matthies	Sparse data formats and efficient numerical methods for uncertainties quantification in numerical aerodynamics
2010-02	D. Grunwald, M. Lochau, E. Börger, U. Goltz	An Abstract State Machine Model for the Generic Java Type System
2010-03	M. Krosche, R. Niekamp	Low-Rank Approximation in Spectral Stochastic Finite Element Method with Solution Space Adaption
2011-01	L. Märtin, M. Schatalov, C. Knieke	Entwicklung und Erweiterung einer Werkzeugkette im Kontext von IT-Ökosystemen
2011-02	B. V. Rosić, A. Litvinenko, O. Pajonk, H. G. Matthies	Direct Bayesian update of polynomial chaos representations
2011-03	H. G. Matthies	White Noise Analysis for Stochastic Partial Differential Equations
2011-04	O. Pajonk, B. Rosić, A. Litvinenko, H. G. Matthies	A Deterministic Filter for non-Gaussian Bayesian Estimation
2011-05	H. G. Matthies	A Hitchhiker's Guide to Mathematical Notation and Definitions